3D Echocardiography Image Compounding

Yao, Cheng

Awarding institution:
King's College London

3D Echocardiography Image Compounding

Cheng Yao

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

Division of Imaging Sciences & Biomedical Engineering

School of Medicine

King's College London

August 2012

Abstract

Echocardiography (echo) is a widely available method to obtain images of the heart, however, echo can suffer due to the presence of artefacts, high noise and a restricted field-of-view. One method to overcome these limitations is to use multiple images, using the "best" parts from each image to produce a higher quality "compounded" image.

This thesis describes a new method to allow multiple 3D echo images to be compounded into a single better quality volume. I have proposed a definition for an "ideal" compounded image and have used this to guide the design of my compounding method, in particular designing a method to reduce the effect of image artefacts and to make use of larger numbers of images. My compounding method has been validated using phantom, volunteer and clinical images. The overall motivations for improving echo image quality are twofold: Firstly to provide clinicians with higher quality images which I hope will improve the accuracy of clinical decision making. Secondly to provide higher quality images for subsequent post-processing algorithms.

A number of methods have been proposed to compound sets of ultrasound images, all of which have reported improvements in image quality. However, previous 3D compounding methods have typically been applied to

a relatively small number of images (most of them only use two images, and only one uses six images). I have investigated the effect of compounding with larger numbers of images. Results showed continued improvement in image quality up to ten images (the maximum number we deemed feasible to acquire in a clinical setting and it is approximately double of images used previously).

Artefacts occur regularly within echo images, particularly shadowing artefacts (due to the highly reflecting interfaces caused by the ribs and lungs when imaging the heart). However, previous 3D compounding methods haven't directly claimed and demonstrated the effect of artefacts. Therefore, I have proposed a 3D compounding algorithm which specifically aims to reduce the effect of echo artefacts (shadowing) as well as improving the signal-to-noise ratio, contrast, and extending the field-of-view. My method to reduce the effect of artefacts is to weight image information from different views based on a local feature coherence/consistency. I hypothesize that the presence of an artefact in an image varies greatly depending on view direction, therefore much lower consistency values will be calculated for artefact regions enabling them to be detected, and their influence on the compounded image to be greatly reduced.

The accuracy of the image registration is important and errors will likely affect the final compounded images quality. In addition to registration accuracy my system needs to work robustly and have a large enough capture range to enable automatic registration from a suitable starting position. Therefore, experiments have been carried out to calculate registration consistency and capture range for the semi-simultaneous registration method

and an improved version later on for a direct comparison.

Compounded image quality validation has been carried out using phantom and volunteer datasets consisting of up to ten multi-view 3D images. Multiple sets of phantom images were acquired, some directly from the phantom surface, and others by imaging through soft and hard tissue mimicking material to degrade the image quality. My compounding method is compared to the original, uncompounded echo images, to two basic statistical compounding methods (Mean and Maximum) and to the state-of-the-art phase based method. Results show my method is able to take a set of ten images, degraded by soft and hard tissue mimicking material, and produce a compounded image of equivalent quality to images acquired directly from the phantom. My method when applied to phantom and volunteer data achieves almost the same signal-to-noise improvement as the mean method, while simultaneously almost achieving the same contrast improvement as the maximum method.

In order to investigate clinical use of my compounding, two types of pathology have been chosen: Transposition of the Great Arteries and Hypoplastic Left Heart Syndrome. Seven datasets in total have been compounded and validation has been carried out by a visual inspection study. Clinicians showed a very strong preference for my compounded volumes in terms of overall image quality, large field-of-view, high endocardial border definition and low cavity noise. Evidences showing the benefit of using compounded echo image for post-processing, are also coming from my collaborative work with Universitat Pompeu Fabra (myocardial motion and strain estimation) and University College London (whole heart segmentation).

Acknowledgements

Firstly I would like to thank my supervisors, Graeme Penney and Tobias Schaeffter, for the initial inspiration and all their help and encouragement. I am very grateful to the rest of the Division of Imaging Sciences & Biomedical Engineering at St Thomas' Hospital, particular thanks go to Andy King and Shahrum Gilani. I would like also thank the clinicians who have shared their expertise and valuable time with me, particular thanks go to John Simpson.

Thanks to Mum for everything she has done for me - the page isn't big enough for that one. A big thank you also goes to Xiaolin for putting up with me and hopefully will continue doing so for a few more years.

Lastly and most importantly, in memory of Grandma - my past, present and future luminary.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Ultrasound is sound with a frequency greater than 2 MHz (frequencies between 2-18 MHz are typically used in diagnostic ultrasound) and it is widely used in clinical diagnosis and evaluation, especially in cardiac imaging where it is referred to as echocardiography or echo. The advantages of using echocardiography compared to other modalities such as magnetic resonance (MR) or x-ray computed tomography (CT) are that it is relatively inexpensive, does not involve ionising radiation, is portable and allows for real-time imaging.

However, echocardiography imaging has certain quality issues, such as low signal-to-noise ratio and low tissue contrast. Image quality depends greatly on both the skills of the operator and patient specific issues. Firstly, echocardiographers develop specialized knowledge and skills enabling them to acquire much better quality images compared to a person new to echo imaging. By modifying the probe position to find good acoustic windows or adjusting machine parameters, information in the acquired image is optimized. Secondly, the image quality is very patient specific. For example,

patients with more intervening tissue between the transducer and target organ would have worse image quality compared to patients with less, usually fat or muscle, layers. This occurs because intervening layers cause attenuation so a weaker signal is returned which typically results in noisier images.

In a typical echocardiography examination, the echocardiographer will view thousands of images in real-time and save only a tiny percentage for review and possible post-processing. Although the real-time, interactive nature of echocardiography imaging provides very useful information, it also places constraints on image quality. A possible method to improve image quality is to trade off the real-time nature of ultrasound imaging, by combining useful image information from variety of acquisition positions, in order to produce a final image with better quality. This process is often referred to as image compounding. The final image has the potential to benefit clinical diagnosis in terms of both better visual inspection and post-processing.

1.1 Aim and clinical motivation

Recent advances in probe design (i.e. matrix array transducers) have led to the development of 3D echocardiography transducers which can acquire a "volume" of the heart, in contrast to a single 2D cross-sectional plane. The facility to reanalyze the 3D volume in any user-defined plane has led to increasing use of 3D echo in clinical practice. However, all echocardiographic techniques are dependent on an adequate "acoustic window" which refers to there being a a suitable soft tissue pathway (e.g. no air or bone)

from the transducer to target organ to allow adequate echo images to be acquired. Also the field-of-view achievable from a single point of insonation is limited. The image is not of uniform quality due to noise in homogenous tissue regions (speckle noise) and the strongest reflection occurs when the ultrasound beam is normal to the interface between tissues (often referred to as the angular dependence of tissue boundary reflections). In the rest of this thesis, unless given a different or specific name, the reader should assume that the word "artefacts" refers to "shadowing artefacts".

The real-time nature of echocardiography places constraints on image quality. For example, noise and artefacts in images make it more difficult to see the anatomical features clearly, or to accurately calculate the standard cardiac parameters such as ejection fraction. Also, Dilated cardiomyopathy (a heart muscle disease with enlarged heart) is the major cause for cardiac transplantation. The adults prevalence is 1 in 2500 (an incidence of 7 per 100000 per year) (Abelmann, 1985; Tavazzi, 1997; Rakar et al., 1997; Cooper, 2002; Towbin et al., 2006; Jefferies and Towbin, 2010). Therefore, for patients with larger hearts, it would be difficult to see a whole heart and the relationship between anatomical structures in a single acquisition. Furthermore, the image quality (Noble and Boukerroui, 2006) and field-of-view (Rajpoot et al., 2009) greatly affect the performance of echo (or ultrasound in general) image post-processing algorithms. The constrained image quality cannot be totally solved by altering probe design, for example, due to the characteristics of ultrasound physics, and the transducer size (which affects field-of-view) is restricted by the small acoustic windows which allow imaging access to the heart (more details will be covered in Section 1.2.1).

The overall aim is to improve echo image quality, which will benefit both clinical diagnosis and post-processing of echo images. This broad aim can be split up into several additional benefits and these aspects are described below:

1. Less operator dependent image acquisition: clinicians will have various levels of skill and/or different preferences when operating the echo machine. The acquired echo images will vary in terms of image quality and/or anatomical coverage. Image compounding should be able to provide more standard and consistent images with better image quality and information content, resulting in an image acquisition process which is less dependent on operator skill and individual preferences.

2. Easier review and more accurate diagnosis: as mentioned above compounding has the potential to produce more standardized, better quality images. This should result in the images being easier for clinicians to directly interpret, and easier for clinicians not directly involved in the image acquisition to review the images. The improved image quality and field-of-view will provide better definition of anatomical structures and heart functionality so that more accurate diagnosis can be carried out.

3. Widening patient population for echo imaging: as stated in points 1 and 2, with less operator dependency and enhanced diagnostic accuracy, compounding could increase the use of echo imaging to a larger patient population, including patients with limited acoustic windows. Therefore, instead of these patients requiring more expensive and time

consuming imaging systems like MR, or modalities which expose patients to ionizing radiation like CT, they can be referred to echo scans. Thus benefiting patients and the health service in terms of a low cost and more effective service.

4. Reducing training requirements: the benefits of compounding to make echo imaging less dependent on the skill of the operators could have knock on effects in terms of training requirements. The ability to more easily produce higher quality images, and less reliance on acquiring specific anatomical planes should enable echocardiographers to reach the required level of skill in a shorter time frame. In addition, the improved ability to re-review images could enable much more experienced clinicians to review images acquired by less experienced clinicians much more effectively. Which could result in early training concentrating on image acquisition rather than requiring both acquisition and diagnosis by the same individual.

5. Improved performance of subsequent image processing: with improved image quality and enhanced information content, the compounded echo images could be used to facilitate novel post-processing. It could also be useful to improve the results of current post-processing techniques on echo images, e.g. segmentation, registration, motion estimation, etc.

6. Wider variety of applications: additional beneficiaries could be for medical ultrasound imaging in general (i.e. not just in the field of echocardiography), such as for liver imaging, for musculoskeletal ima-

ging, etc.

In the next section, I propose my specific solution, explain image compounding and how compounding techniques can improve echo image quality and field-of-view. Related image quality enhancement techniques are also introduced and I discuss why these are not suitable for my application.

1.2 Compounding techniques

As discussed in the previous section, echo has disadvantages which result in constrained image quality and limited field-of-view. Compounding techniques overcome the problem of different parts of the heart being more clearly visualized when being insonated from a particular position. In normal clinical practice, an echocardiographic study includes multiple images taken from different positions on the chest. These images are, however, reviewed separately from each other in a conventional cross-sectional echo study. Even using current 3D echo techniques, acquired 3D volumes have to be analyzed separately. My technique aims to merge the "most useful" information from a number of 3D echocardiographic volumes taken from different probe positions into a single compounded image, which has improved image quality. In addition compounded images are able to extend the field-of-view which can aid understanding of the spatial relationship between anatomical structures which could not be well visualized from a single projection.

Whatmough et al. (2007) has defined image compounding as as follows: "Image compounding is a method for improving sonographic image quality

whereby several image frames are combined in an attempt to cancel out random variations (noise) and hence increase the signal-to-noise ratio."

For the work described in this thesis I particularly wish to improve echo image quality. As such I aim to reduce the angular dependence of echo images, reduce speckle and reduce artefacts. I propose that all of these effects which reduce image quality, can be more effectively countered if multiple different views of the same anatomy are combined. I therefore extended the definition proposed by Whatmough et al. (2007) to involve multiple view directions.

Therefore, my definition of image compounding is: the use of multiple images of the same anatomy from different view directions, to produce an output image with improved quality. Specifically my compounding will require:

- multiple images of the same object
- images taken from multiple view directions
- known spatial relationships between images

My compounding process will determine the best representative image intensity to represent an imaged object, given a set of images which provide multiple intensity values for the object at the same spatial location.

Figure 1.1 shows an example of how my image compounding will work. Echocardiography images are acquired from different view positions, some parts of the heart are imaged well in some images but not in other images. Image alignment is used to match all the images together into a common co-ordinate system. For each corresponding location, multiple intensity values

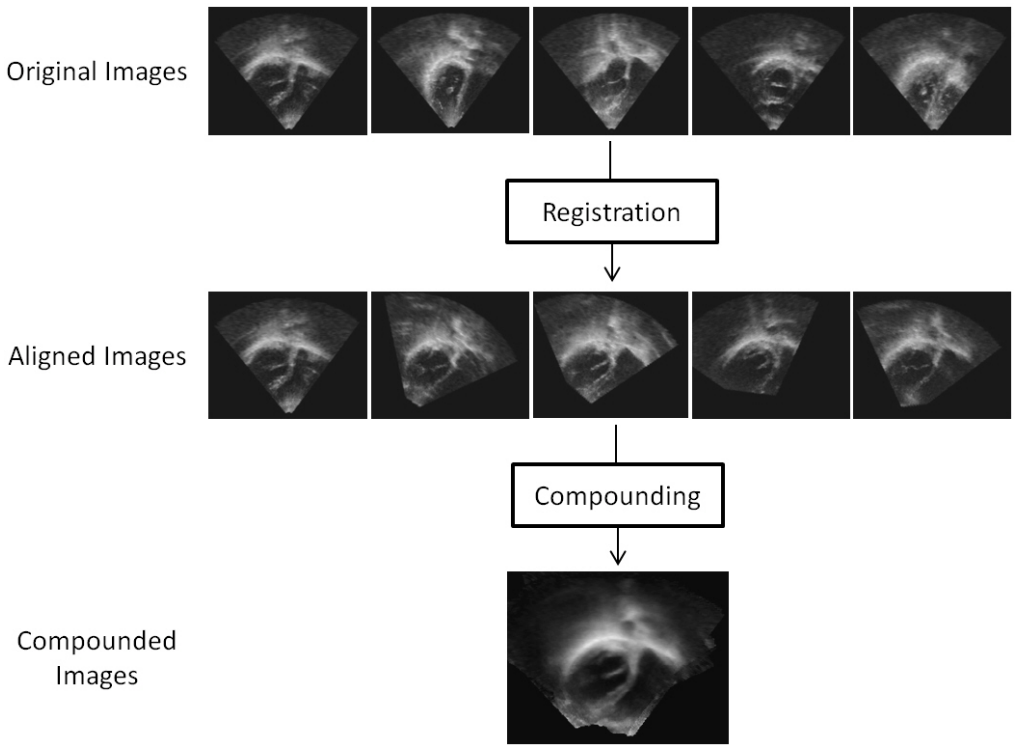


Figure 1.1: Diagram to show an example of compounding. The upper row shows images are taken from different view angles from the same heart; Middle row shows those images are aligned in the same coordinate system, which means corresponding anatomical structures are matched to each other in same location; Bottom image shows the compounded image. The compounding process has taken the multiple intensity values for a given anatomical location from the aligned image (middle row) and produced a single intensity value for each position in the final compounded image.

are collected, and my compounding process determines the best intensity value to be used in the final compounded image at that location.

The potential benefits for improving echo image quality are twofold: Firstly to provide clinicians with higher quality images which will improve the accuracy of clinical decision making. Secondly to provide higher quality images for subsequent post processing algorithms, e.g. ventricle segmentation and wall motion estimation for calculation of ejection fraction, or to analyze dyssynchrony.

1.2.1 Related image enhancement techniques

I now briefly review a few related techniques in literature designed to improve image quality:

Transducer based compounding There has been a form of compounding within a B-mode image which has been available commercially for a number of years. This involves using electronic beam steering by the transducer to acquire images of a given region of tissue from multiple directions without moving the probe. The resultant images are of higher quality primarily due to reduced speckle. It has been rapidly adopted for use in breast imaging (Forsberg, 2004), and has found applications in peripheral vasculature and musculoskeletal imaging (Entrekin et al., 2001). However effective compounding requires sufficient angular differences between view directions (Figure 1.2 left) and maximum angular differences decrease as the size of the transducer footprint is reduced and as the depth of the target tissue increases (hence the main applications are in peripheral tissue where large linear arrays can be used). Therefore, for cardiac imaging, where the transducer footprint needs to be small (due to small acoustic windows), the target tissue is deep within the body and a high frame rate is often crucial, transducer-based compounding is not feasible (Figure 1.2 right).

Temporal compounding Perperidis et al. (2009) use the term temporal compounding to refer to a method which compounds multiple images from the same view angle to reduce random noise. However I am

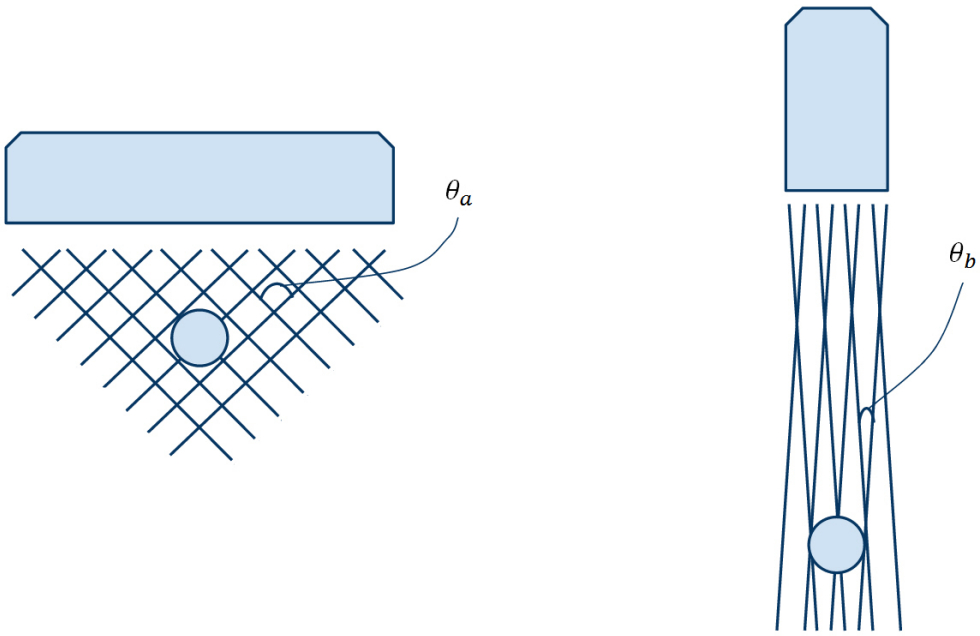


Figure 1.2: Diagrams to illustrate transducer based compounding. The left diagram demonstrates an ultrasound transducer with a wide footprint. The object to be imaged is close to transducer so a large angular difference θ_a is obtained for effective compounding. However, for cardiac imaging, a small transducer footprint is needed (above right) and the target object is deep within the body. This results much smaller angular differences θ_b so is much less effective for compounding.

especially aiming to reduce angular dependent effects and artefacts, therefore compounding images from multiple view directions is more appropriate for my application.

Mosaicing The term mosaicing (Wachinger et al., 2007) is often used when the images to be combined have little overlap and the primary interest is to extend the field-of-view. For my application, the main objective is increasing the image quality (such as improving signal to noise ratio and contrast), so input images with substantial overlap are required.

Previous compounding methods More details can be found in Chapter 3. Previous methods (Leotta and Martin, 1999; Grau and Noble, 2005; Soler et al., 2005a; Rajpoot et al., 2009; Szmigielski et al., 2010) have all shown image quality improvement. However, there are several aspects which can be improved: Firstly, they have typically been applied to a relatively small number of images (most of them use two images, one uses six). The use of more image data should average out random noise as given by statistics theory. Secondly, artefacts occur regularly within echo images, particularly shadowing artefacts (due to highly reflecting interfaces caused by the ribs and lungs when imaging the heart), but none of the previous compounding methods has specifically addressed the effect of echo artefacts.

1.3 Objectives

The main objectives of my PhD work are presented below:

Investigating the effect of compounding larger numbers of 3D

echo images on image quality. Although statistical theory suggests that more images should improve image quality, other factors may arise during the compounding process that may decrease image quality as more images are used. For example, when more images are used, registration errors might cause anatomical structures to be blurred. Therefore, an investigation will be carried out to show the quality improvement effect on larger numbers of 3D data. This will be carried out on a relatively fast moving anatomical structure (the heart) where both temporal and spatial registration errors will arise.

Defining an "ideal" compounded image. This thesis presents, for the first time in literature (to my knowledge) a definition for the target results of my compounding algorithm. I begin by proposing an "ideal" compounded image. Then I describe how my methodology attempts to produce my proposed "ideal" image. Other possible definitions of an "ideal" image are also discussed.

Proposing a novel compounding method based on feature consistency. My proposed method differs from previous 3D compounding techniques as it has been specifically designed to reduce the effects of artefacts, in addition to the standard cited reasons for compounding, i.e. improving signal-to-noise ratio and contrast, and extending the field-of-view. My method to reduce the effect of artefacts is to weight image information from different views based on a local feature coherence/consistency. I hypothesize that the presence of an artefact in an image varies greatly depending on view direction, therefore much

lower consistency values will be calculated for artefact regions enabling them to be detected, and their influence on the compounded image to be greatly reduced.

Creating a validation method that allows much more clinically realistic phantom images to be acquired. My solution is to use multiple sets of phantom images, some acquired directly from the phantom surface to produce a "gold-standard", and the others acquired by imaging through hard and soft tissue mimicking material to degrade the image quality to make images more clinically realistic. Numerical values are calculated which define some factors of image quality and results show how compounding can reduce image noise and artefacts.

Volunteers study. Ten volunteer datasets are used to test the performance of my feature consistency based compounding algorithm. Both numerical measurements and visual inspection have been carried out.

Clinical evaluation. Seven patient datasets from two specific clinical pathologies: Transposition of the Great Arteries (TGA) and Hypoplastic Left Heart Syndrome (HLHS) are used to test the performance of my feature consistency based compounding algorithm. Visual inspection has been carried out. When moving from volunteer to patient data, there are issues involved with respect to registration algorithm performance.

1.4 Overview of thesis

The main sections in this thesis are defined as below:

Clinical and technical background : Chapter 2 provides a tailored background to show the aspects of ultrasound physics and cardiac ultrasound imaging, which are relevant to my research. The issues involved in echo imaging are discussed (e.g. angle dependent boundary reflection, non-uniform intensity of homogenous tissue, and artefacts). Then in Chapter 3 a solution is proposed to overcome echo image quality constraints by using compounding techniques, and a comprehensive review of previous compounding techniques is presented.

3D echo compounding using more images : An investigation has been carried out on larger numbers of images and the results are shown on the performance of compounding more 3D echo images in Chapter 4.

Consistency based compounding method : In Chapter 5 I define what an "ideal" compounded image should look like and propose a 3D compounding method which has been specifically designed to reduce the effects of artefacts.

Validation : Thorough validation has been carried out on a realistic heart phantom and volunteer data in Chapter 6. The use of my compounding technique on two specific clinical applications is described in Chapter 7.

Conclusion : At the end of this thesis, Chapter 8, I draw a conclusion to this thesis and discuss potential investigations for future work.

Chapter 2

Background

Ultrasound is extensively used in clinical diagnosis. In this chapter, I begin with a brief general background of ultrasound and echo image formation. This covers ultrasound physics and how ultrasound waves interact with human tissue. Then an overview of common echo artefacts, which is followed by the use of echo for cardiac examinations and how clinicians can alter imaging parameters to obtain an optimum image.

2.1 Ultrasound image formation

2.1.1 Basic ultrasound wave physics

Ultrasound has a frequency (ν) greater than 20,000 Hz, which is out of the human hearing range (20 to 20,000 Hz). Ultrasound takes the form of longitudinal waves (Figure 2.1). The waves oscillate in the direction of wave propagation, i.e. the medium vibrates in the same or opposite direction of wave motion, consisting of high and low pressure disturbances.

The wavelength is the length of one cycle of the wave, measured parallel

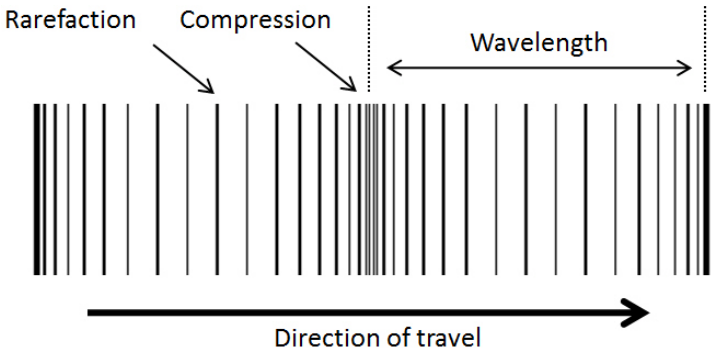


Figure 2.1: In a longitudinal wave, particle motion is aligned with the direction of travel, resulting in bands of high and low pressure (compression and rarefaction respectively).

to the direction of travel (shown in Figure 2.1). It is related to speed (c) and frequency (ν) by:

$$\lambda = \frac{c}{\nu} \tag{2.1}$$

The speed of ultrasound is determined by the density (ρ) and stiffness (κ) of the medium it is travelling in. The speed of sound (c) is then described as:

$$c = \sqrt{\frac{\kappa}{\rho}} \tag{2.2}$$

The acoustic impedance (z) is a measure of the ease of propagation, and is determined by medium's density and stiffness.

$$z = \sqrt{\rho\kappa} = \rho c \tag{2.3}$$

When a wave travels through a boundary of two mediums with different acoustic impedance, some of the energy is reflected back towards the source of the wave and the remainder is transmitted into the second medium. Accordingly, they are called reflection and refraction (Dendy and Heaton, 1999) (Figure 2.2).

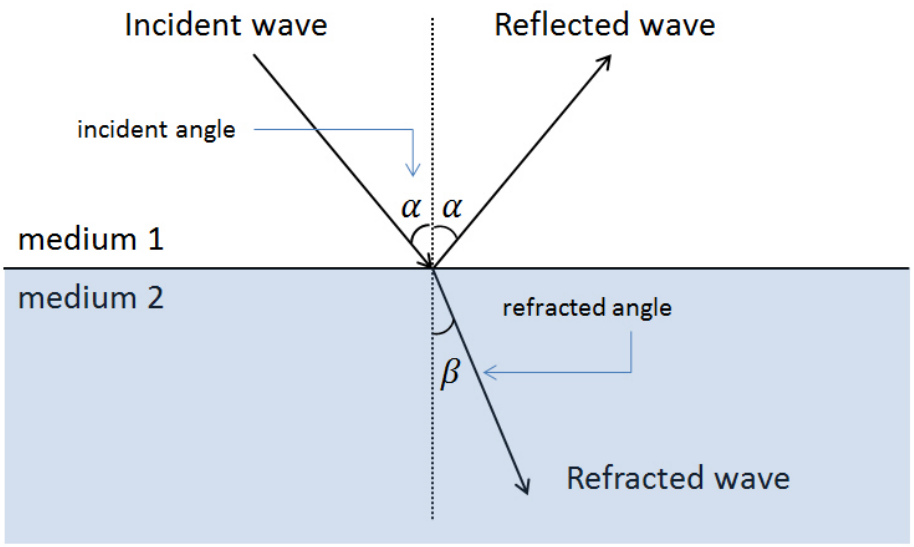


Figure 2.2: Ultrasound wave reflection and refraction on an interface of two mediums with different acoustic impedance.

When a reflected wave is produced, it carries a fraction of the power of the original wave. If the interface is large, smooth, and flat, and the incident wave is normal to the interface, intensity reflection coefficient R can be defined as:

$$R = \left[\frac{z_2 - z_1}{z_2 + z_1}\right]^2 \tag{2.4}$$

Its value depends only on z_1 and z_2: the characteristic acoustic impedances of the medium on the incident and transmission sides of the interface respectively. If the incident angle is not normal to the interface, only a portion of the wave will be reflected, depending on the angle between interface normal and propagation direction of the wave. The remaining refracted wave changes speed and direction of propagation as it crosses the interface of two mediums.

When the wave is incident on very small-scale targets (of size comparable to or less than the wavelength), the wave is scattered over a large range of angles. This is known as scattering, shown in Figure 2.3. For a target

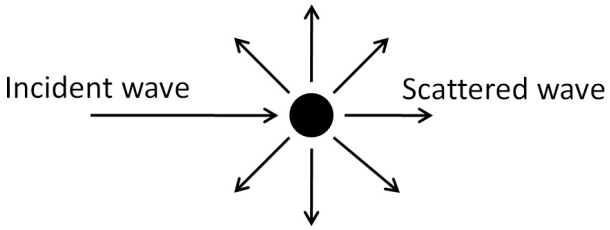


Figure 2.3: Ultrasound wave scattering on an small-scale object.

which is much smaller when compared to the wavelength, the wave may be scattered uniformly in all directions.

When an ultrasound wave propagates though a medium, the energy of the wave decreases with the distance travelled. The effect is known as attenuation. The effect of attenuation shall be described with attenuation coefficient (δ, which is normally expressed in units of $dB\,cm^{-1}$). δ is dependent upon the frequency of the ultrasound waves and δ/ν remains approximately constant. In other words, higher frequency ultrasound will experience higher attenuation when travelling through a medium. The measured intensity $I_{measured}$ of an ultrasound wave after travelling through a layer of material with thickness d is related to the incident intensity $I_{incident}$ according to the inverse exponential power law:

$$I_{measured} = I_{incident} e^{-d\delta} \tag{2.5}$$

2.1.2 The propagation of ultrasound waves in tissue

The human body consists of many different tissue types and structures, such as blood, muscle, bone, fat and air. When an ultrasound wave propagates within tissue, various effects, such as reflection, refraction, scattering and attenuation can occur.

The frequency range used in medical diagnostic ultrasound are between

Table 2.1: Ultrasound properties (approximate) of some human tissues at 37°C, and other media (Duck, 1990; Kaye and Laby, 1995).

Medium	c (ms^{-1})	z $(kg\,m^{-2}s^{-1})$
Blood (whole)	1585	1.68×10^6
Liver	1580	1.66×10^6
Muscle (skeletal)	1575	1.64×10^6
Fat (average)	1430	1.31×10^6
Bone (skull, outer)	2800	5.60×10^6
Water (20°C, 1 atm)	1482	1.48×10^6
Air (20°C, 10% humidity)	311	392

2 to 18 MHz. However, tissue differs much more in stiffness than in density, so that bone (very incompressible, high stiffness) has a higher speed of sound than muscle, despite the fact that it is more dense. The mean value of the speed of sound in soft tissue is typically considered to be 1540 ms^{-1}. Table 2.1 presents ultrasound properties of some human tissues, which includes the speed of sound and acoustic impedance in specific tissues.

The value of the of intensity reflection coefficient, R, can be calculated using the characteristic acoustic impedances of two different mediums as described in equation 2.4. In general, the typical value of R between different human soft tissues is small, e.g. with 1.45×10^{-4} of R at muscle and blood interface. However, boundaries like air and bone have significant differences in acoustic impedances compared to other tissue mediums, which will result in strong reflections, e.g. with an R of nearly 1 at the interface between air and other tissue types, with an R of around 0.6 between Bone and other tissue types.

When an ultrasound wave encounters an object which is smaller than its wavelength, it generates scattered waves that radiate away in all directions.

Table 2.2: The approximate values of attenuation coefficient for some human tissues (Dendy and Heaton, 1999).

Medium	δ/ν $(dB\,cm^{-1}MHz^{-1})$
Blood	0.2
Liver	0.6
Muscle	1
Bone	22
Water	0.0022

Speckle patterns will then be produced by constructive and destructive interference of those scattered waves, depending on different tissue types or different pathologies.

Table 2.2 shows approximate values of attenuation coefficient for some human tissues. Values typically range from 0.6 to 1 $dB\,cm^{-1}MHz^{-1}$ for soft tissues (e.g. muscle) and lower for blood.

2.1.3 Production of images using ultrasound

A typical ultrasound imaging system consists of the following three main components: Transducer, Signal Processor and Image Display.

In ultrasound imaging, the transducer is placed in direct contact with the patient's body. In order to avoid large reflection occurring at the transducer-skin interface a coupling gel is routinely used. The transducer produces ultrasound pulses and receives the returning echoes. The actual sound generating and detecting component is a thin piezoelectric plate. The piezoelectric material converts an electrical current to pressure (the pulse wave), and, conversely, generates electrical current when subjected to mechanical stress (returning echoes). A transducer contains multiple piezoelectric elements within a small transducer area. The arrangement of multiple piezoelectric

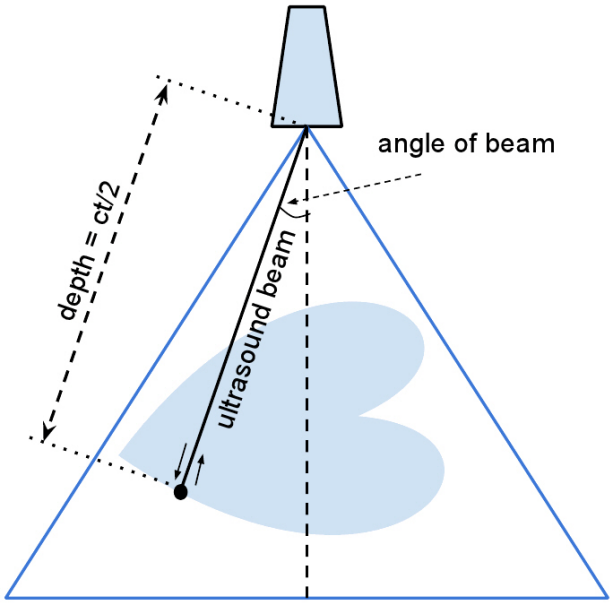


Figure 2.4: The position of the reflector/scatterer is determined by the angle of beam and the depth. t is the measured time for the transmitted signal to return back to the transducer. The depth of the reflector/scatterer is calculated as $ct/2$ where c is the estimated average speed of sound within the tissue. The echo machine sweeps the beam across the image sector, therefore the beam angle is known. Knowledge of angle and depth enable the reflector/scatterer position to be determined and an image to be produced.

elements can be in a straight line (1D array) or form a 2D matrix (2D matrix array), which allows 2D or 3D imaging accordingly.

Whenever a ultrasound pulse encounters a interface with different acoustic impedances or homogeneous tissue with small scale variations in acoustic properties which causes scattering, part of the sound wave is reflected back to the probe and is detected as a returning echo. The position of the reflector/scatterer is then determined by the angle of beam and the depth. As Figure 2.4 shows, the angle of beam is know by the beam formed from the transducer and the time which takes for the echo to travel back to the probe is measured as t. Therefore the depth can be calculated by multiplying $t/2$ by the speed of sound.

The amplitude of the transducer generated ultrasound signal are typically very small, so it needs to be amplified before manipulated and displayed directly. The amplifier is used to amplify all ultrasound signals equally (linear amplifier). This is called overall gain and can be adjusted by the user of the ultrasound system (more details will be covered in Section 2.2.3). The effect on the output image is to make all image intensities brighter or darker, whatever their depth in the image.

As discussed in Section 2.1.1, when an ultrasound wave propagates through tissue and returns back to the transducer, the signals are attenuated, i.e. the returning signal from a large depth in tissue is much weaker than that from closer one (if a similar interface is considered). Therefore, the signals from deep tissues need to be amplified in order to compensate for its attenuation. This technique is commonly called time-gain compensation (TGC). The TGC is usually controlled automatically by the ultrasound system and is determined by the interface depth in tissue and ultrasound frequency. However the ultrasound machine cannot identify tissue types where the attenuation coefficient is different, so TGC can also be manually adjusted by the user of the ultrasound system (more details will be covered in Section 2.2.3).

When an ultrasound waves is incident on an interface or a scatterer, the range of amplitudes detected from returning echoes of different targets is very large. For example, reflected intensity ranges from less than 1 % of the incident intensity for a tissue-tissue interface to almost 100 % for a tissue-air interface, and the intensities of returning echoes from small scatterers are usually much smaller than those from large interfaces (Hoskins et al., 2003).

To be useful diagnostically, a output image should contain both echoes from typical interfaces and tissue scattering. Therefore, a compression technique (non-linear amplifier) is used, and weak echoes from scattering within the tissue are boosted more than the large echoes from interfaces, so that they can be effectively displayed at the same time.

The above paragraphs have described how an ultrasound image is formed. When forming an image, the ultrasound system makes a number of assumptions about ultrasound propagation in tissue. The key assumptions are (Bigatello et al., 2000; Frankel and deBoisblanc, 2010; Klein and Asher, 2011):

- sound travels in a straight line
- sound travels to target reflector/scatterer directly and comes back to the transducer directly
- the speed of sound is constant and known
- The amplitude of returning waves is related directly to the reflecting or scattering properties of distant objects

As described in Section 2.1.1 and 2.1.2, the speed of sound and attenuation coefficient varies with tissue type, and sound can be refracted at interfaces. These effects will cause violations of the above assumptions. Significant variations from these conditions are likely to cause image artefacts, more details of which will be covered in the next section.

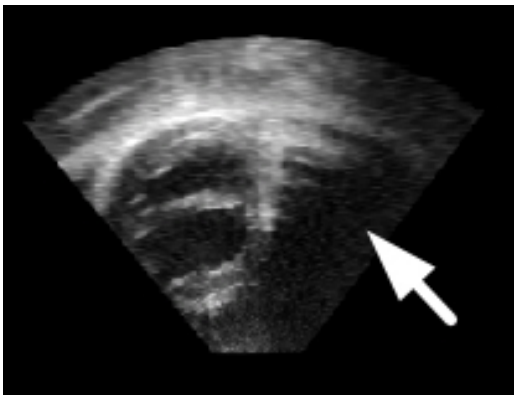


Figure 2.5: Example of shadowing artefact.It reduces the signal along the beam direction (region highlighted by arrow).

2.1.4 Artefacts

Artefacts commonly exist in ultrasound images. Artefacts can be defined as missing information or extra misleading information appearing in an image. They are typically caused by the physical processes of the ultrasound beam and violate the assumptions about the ultrasound wave that are used to produce an image as described in the previous section. Although some of these artefacts may actually provide useful information, the majority may confuse the clinicians. The most commonly encountered artefacts in echo imaging are outlined below.

Shadowing is caused by diminished sound or loss of sound behind a strongly reflecting structure (e.g. large calcification, bone) or strongly attenuating structure (e.g. solid tissue, significantly dense or malignant masses). In echo imaging, it is typically caused by poor probe-patient contact or ribs blocking the sound wave. Figure 2.5 shows an example of shadowing.

Enhancement where an abnormally high brightness is seen. This is caused when sound travels through a medium which has a lower attenuation

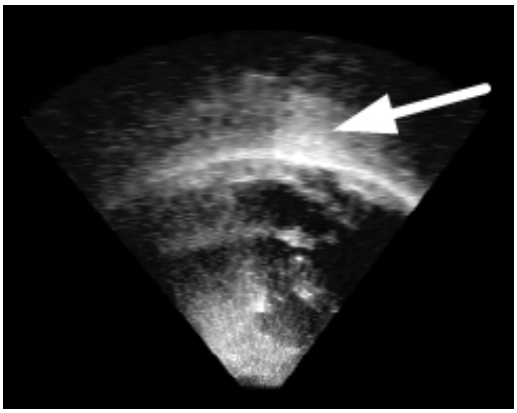


Figure 2.6: Example of enhancement artefact is shown as high brightness reflection (region highlighted by arrow).

coefficient than the surrounding tissue. The applied TGC does not match the actual attenuation rate in the target tissue. For example, in Table 2.2, the attenuation of the sound in blood is less than that of muscle (the surrounding tissue). Reflectors at position deeper than the low attenuation region are much brighter compared to neighboring tissues where the signal has been subjected to higher attenuation. See Figure 2.6 for an example.

Reverberation Reverberation artefacts appear as multiple spaced lines perpendicular to the direction of the ultrasound beam. Reverberation is caused by the sound reflecting back and forth between tissue interfaces and then back to the transducer, which is a violation of the assumption on sound travelling directly to target and back to the transducer. See Figure 2.7 for an example.

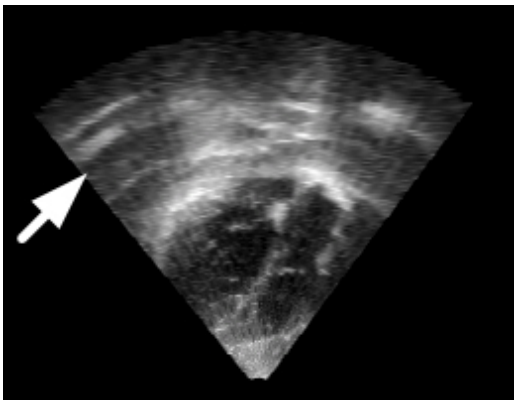


Figure 2.7: Example of reverberation artefact is shown as parallel lines (region highlighted by arrow).

2.2 Heart image acquisition using echo

In this section, a brief description of the structure and function of the heart will be presented, and the routine clinical methods for echo imaging of the heart will be described.

2.2.1 Structure and function of heart

Figure 2.8 shows the four chambers of heart. The atria are the upper chambers and the ventricles are the lower chambers. The thick interventricular septum separates the left and right side of the heart. Myocardium is the thick tissue. Endocardium is a thin membrane which lines the chambers and prevents the blood from clotting. Valves separate the atrium and ventricle on each side to avoid the blood flowing back. The pericardium is membrane to protect the heart muscle and also to hold the heart in its place in the chest.

The function of heart is to keep the blood flowing throughout the body and supply the other organs. A heartbeat is a cycle of systole (atria-ventricle contraction) to diastole (atrial-ventricle relaxation). The contraction of

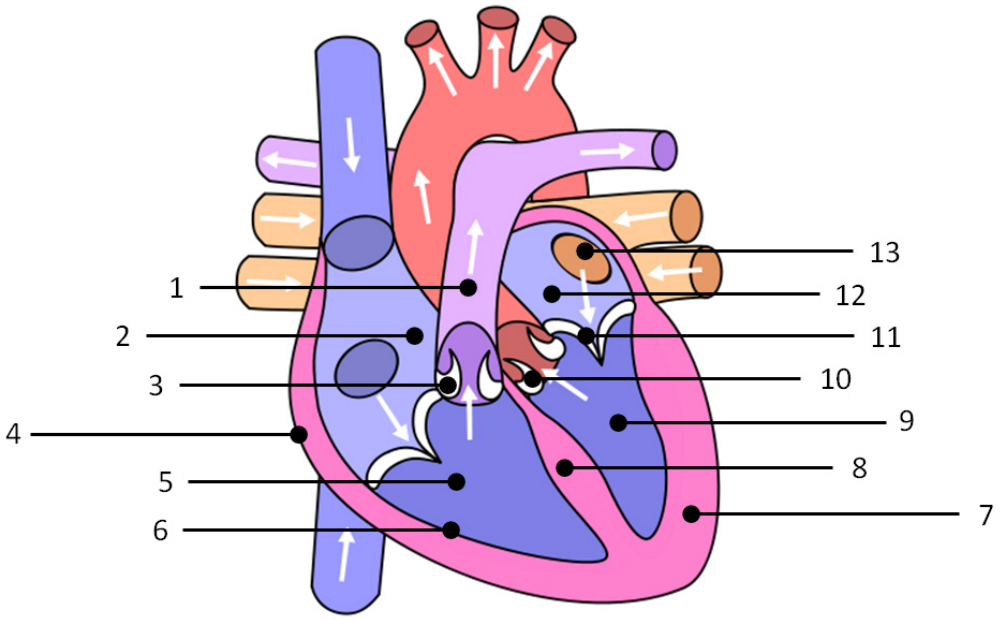


Figure 2.8: Diagram of human heart structure and showing direction of blood flow. (1) pulmonary artery; (2) right atrium; (3) pulmonary semilunar valve; (4) pericardium; (5) right ventricle; (6) endocardium; (7) myocardium; (8) septum; (9) left ventricle; (10) aortic semilunar valve; (11) mitral valve; (13) pulmonary vein. The diagram is adapted originally from http://en.wikibooks.org/wiki/GCSE_Science/The_Heart

right ventricle pushes blood into the pulmonary artery through the pulmonary semilunar valve. It carries deoxygenated blood to the lungs for oxygenation. The blood then flows though the pulmonary vein into the left atrium. The left atrium and left ventricle are separated by the mitral valve. The contraction of the left ventricle pushes blood to the body through the aortic semilunar valve.

2.2.2 Acquisition protocol for echo imaging

Echo imaging is not a easy task. Producing a good image may require appropriate patient position, breath phase or specific view angles. In order to bring the heart closer to the chest wall for better data acquisition, patients are typically positioned in the left lateral decubitus position, or images are taken at end expiration. Clinicians are trained to use the three standard transducer positions for data acquisition (see Figure 2.9), and an example of standard view images are presented in Figure 2.10 accordingly. The protocols of acquiring the standard views are:

Parasternal view is obtained by placing the transducer at the left parasternal area in the third or fourth intercostal space.

Subcostal view is obtained when patients are in the supine position. Relaxing the abdominal muscles by flexing the patient's knees and forced inspiration frequently improve the views.

Apical view is obtained with the transducer usually placed slightly lateral and inferior to the nipple, but it may be substantially displaced laterally and inferiorly because of cardiac enlargement or rotation or

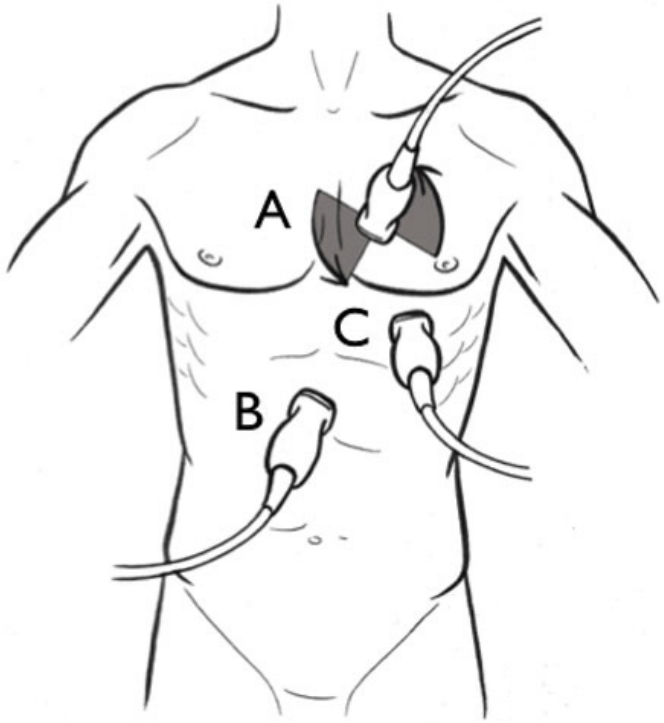


Figure 2.9: Diagram of standard transducer positions. (A) parasternal; (B) subcostal; (C) apical. The diagram is adapted originally from http://67.131.133.95/html/pre/ultr/ultr/041080025.asp

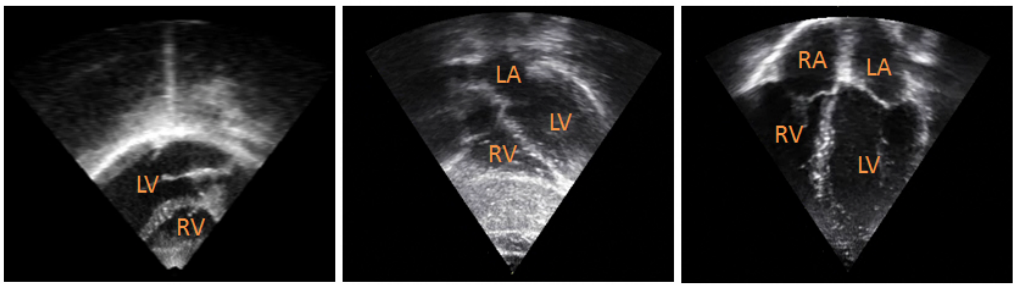
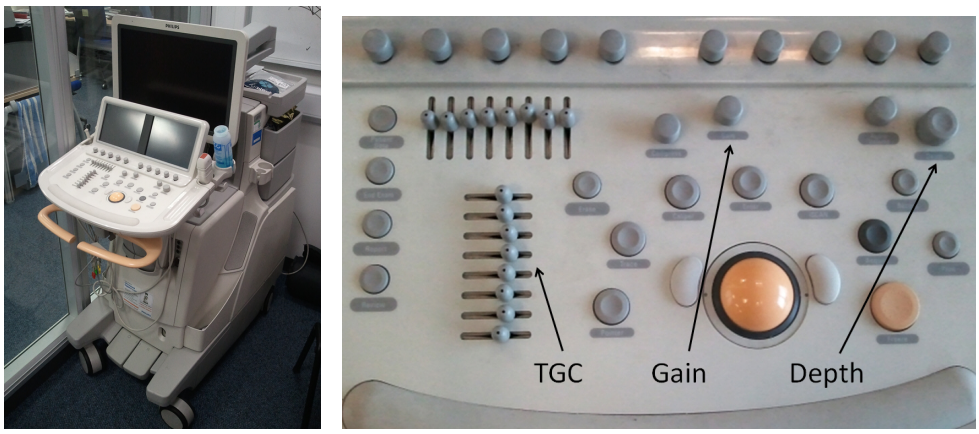


Figure 2.10: Example of standard view images. From left to right, they are parasternal, subcostal and apical. (LV = left ventricle; LA = left atrium; RV = right ventricle; RA = right atrium)



(a) (b)

Figure 2.11: (a) Philips iE33 ultrasound machine; (b) control panel of this ultrasound machine.

both.

2.2.3 Settings on the echo machine

Apart from using good acoustic windows and patient positioning to improve the quality of echo images, it is also possible for the operator to alter a number of settings on the echo machine (Figure 2.11). The main settings are:

Overall gain : This controls the amplitude of returning signals before display. Adding gain could increase the brightness of an image, but over gaining will decrease the image contrast.

Time-gain compensation (TGC) : TGC compensates the energy loss of the ultrasound wave due to attenuation. For example, suppression needs to be applied to the signal from close to the transducer and higher amplification is required for the signal from the deeper depths (the equalizer-like controls in Figure 2.11).

Depth : This control adjusts the overall image depth. A deeper field-of-

view is required for some structures to be imaged from certain cardiac views (e.g. subcostal view).

Acquiring an Electrocardiogram (ECG) is also important for cardiac data acquisition. This measures the electrical signal released by the pumping heart muscle using small electrodes placed on the chest. ECG is used to gate echo image acquisition, so the cardiac phases from different acquisitions can be correlated. ECG can also be used to filter out data with an irregular heart rhythm.

For the 4D echo imaging, there is an inverse relation between frame rate (temporal resolution), volume size (field-of-view) and spatial resolution as discussed in (Badano et al., 2010), which means in order to achieve the maximum frame rate, the volume size and spatial resolution need to be compromised.

There are many more new features and developments moving into clinical practice in cardiac echo imaging. Examples are: volume reconstruction, segmentation, chamber quantification (Hung et al., 2007), and automated qualification of blood flow using Doppler techniques (Thavendiranathan et al., 2012).

2.2.4 Clinical use of 3D echo

Three-dimensional echocardiography (3D echo) provides the improved and expended diagnostic capabilities to the cardiac ultrasound. The early approaches to 3D echo were based on the reconstruction from a set of 2D images and interpolation of missing data (Hung et al., 2007). However, this method has technical limitations in acquisition and offline data processing

is required. The development of 3D transducers and computer technology circumvents many of the disadvantages of the reconstructive method, so that now real-time 3D echo images can be produced. Current real-time 3D echo systems use a 2D matrix array transducer with typically 3000 piezoelectric elements which can acquire cardiac structures in a pyramidal image (Oh et al., 2007). a wide field-of-view 3D echo image can be produced by merging volumes obtained from several cardiac cycles (four to seven) using ECG gating.

The clinical usefulness of echo images can be considered under two categories (Whiting and Wiens, 1977; Lang et al., 2009): (a) the measurement of anatomical structures and their changing relationships throughout the cardiac cycle; (b) motion dependent function abnormalities demonstration. Echo provides clinicians with views of cardiac structure and function. It is very useful for diagnosis and treatment planning. For example, quantification of chambers and wall motion assessment. Quantification of chambers is used in the assessment of valvular heart disease and congenital heart disease, by measurement of left ventricular mass, volumes, ejection fraction and wall motion assessment.

2.3 Conclusion

In this chapter the fundamentals of ultrasound physics and ultrasound image formation have been introduced. Image features like boundary reflection, homogenous tissue region and image artefacts were discussed. Because of the location of the heart in the human body, and the nature of echo imaging,

the image quality and field-of-view are restricted by various factors. It can be challenging for the clinician to acquire useful images, particularly in patients with poor acoustic windows. Therefore, methods to improve echo image quality will be presented in the next chapters.

Chapter 3

Review of echo compounding techniques

In this chapter, prior work on echo compounding techniques is reviewed. Firstly, I define the notations used in the rest of this thesis, and then the different compounding methods in literature are presented and discussed. As image alignment is important for accurate compounding, the registration strategies used in previous compounding papers are introduced. Then the methods used to validate these compounding techniques are discussed.

3.1 Notations

As described in Section 1.2, compounding is a technique which uses multiple images of the same anatomy from different view directions, to produce an output image with improved quality. For the convenience of describing and understanding the echo compounding techniques later on, I define the following notations:

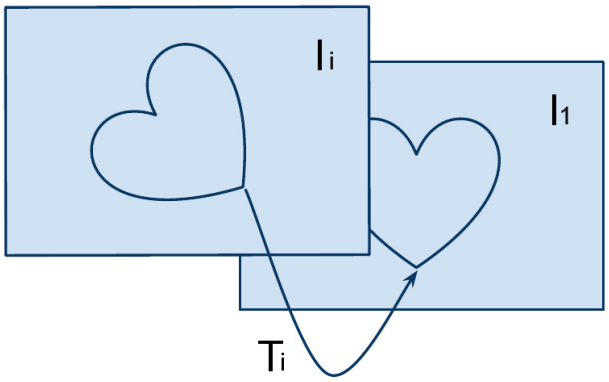


Figure 3.1: Diagram to show how coordinate systems are matched together. The first image I_1 in a dataset is chosen as the reference. T_i is the transformation matrix to transform the image information on I_i back to the reference coordinate system (I_1).

- Images and image intensities: Assuming there are N images in a dataset, I_i represents the i^{th} image ($i = 1, ..., N$) and I_f represents the final compounded image. $I_i(\mathbf{x})$ represents the intensity values at the voxel position $\mathbf{x} = \{x_1, x_2, x_3\}$.

- Coordinate matching between images: Considering images acquired from different view directions, the matrix T_i is used to denote the transformation from the coordinate of image I_i back to reference coordinate system (see Figure 3.1, the coordinate system of the first image in a dataset is chosen as the reference, so T_1 is the identity matrix). For a given voxel position $\mathbf{x}$ in the reference image, $I_i(T_i^{-1}(\mathbf{x}))$ ($i = 2, ..., N$) are the set of corresponding intensities in the other $N-1$ images.

- To simplify notation, $I'_i = I_i(T_i^{-1}(\mathbf{x}))$ is used. These N overlapping intensity values $\{I'_1, ..., I'_N\}$ are used to calculate the intensity value in the final volume I_f using a chosen compounding method.

3.2 Image compounding

There are a number of echo image compounding techniques in the literature. These techniques will be described in this section under the following categories: general image enhancement methods based on statistics, advanced methods and ultrasound specific methods. In this section I assume that all the images have already been accurately aligned.

3.2.1 Standard image enhancement methods

Basic statistical image compounding algorithms are very general and can be used to enhance sets of images from a number of modalities.

In the case where multiple intensity values $\{I'_1, ..., I'_N\}$ are collected for each image voxel position $\mathbf{x}$, statistical methods are used to to summarize the collection of values and to calculate the voxel value in the final volume.

In the literature, echo images are recognized as noisy; this noise can be reduced by using basic statistical image quality enhancement techniques. For example, the mean method (Leotta and Martin, 1999) averages all the overlapping intensities to reduce the effect of randomly distributed noise, however, a blurred output could be produced if artefacts are present or if registration errors exist.

$$I_f = \frac{1}{N}[I'_1 +, ..., +I'_N] \tag{3.1}$$

In order to improve the contrast of an echo image, the maximum method (Leotta and Martin, 1999) ensures that the brightest intensity is passed to the final image. This is useful for reducing some low intensity artefacts,

however, it will also pass high intensity noise values into the final image.

$$I_f = max\{I'_1, ..., I'_N\} \tag{3.2}$$

To keep the best spatial resolution on the features only if they appear in all images, the minimum method can be used to maintain the feature definition, however, it decrease the overall intensity in the final image.

$$I_f = min\{I'_1, ..., I'_N\} \tag{3.3}$$

3.2.2 More advanced methods

Generalized Averaging

The Generalized Averaging method (Soler et al., 2005a) was applied to compound two images to a single volume. The method proposes that an appropriate different operator should be applied to each of these three different scenarios:

- Maximum, to keep a feature if it is present in only one image;
- Mean, to reduce the noise if both images contain uncorrelated speckle patterns;
- Minimum, to improve the definition of feature if it is present in both images.

The authors provide a way to smoothly switch from minimum to average and to maximum operators to utilise image information using the following equation:

$$I_f = \left(\frac{1}{2}((I'_1)^\beta + (I'_2)^\beta)\right)^{\frac{1}{\beta}} \tag{3.4}$$

where $\beta(\mathbf{x})$ is a function of data disagreement, which is computed as:

$$\beta = \gamma \left(|\bar{I}'_1 - \bar{I}'_2| - corr(I'_1, I'_2) \right) \tag{3.5}$$

where $corr(I'_1, I'_2)$ represents the correlation operator on a neighbourhood around $\mathbf{x}$, and γ is used to accommodate the range of the generalized average operator (minimum, mean and maximum).

Soler et al. (2005a) has shown this technique improved the SNR and wall contrast of the myocardium. However, the scenario and algorithm were designed for compounding two images, and the β calculation is difficult to extended to more than two images which could also vary depending on the scenario chosen.

Multiview Deconvolution

Multiview Deconvolution (Soler et al., 2005a) produces a compounded volume by given a point spread function of each acquisition to solve a inverse problem. The authors assumed that each I'_i is a version of the I_f degraded by a point spread function h'_i, which is

$$I'_i = h'_i * I_f \tag{3.6}$$

where h'_i is defined as a Gaussian, with variance depending on the distance to the probe z'_i. Therefore, to find the final volume that best estimates the acquisitions, they define a term

$$Q = \sum_i \parallel I'_i - h'_i * I_f \parallel^2 \tag{3.7}$$

to account for differences between the acquisitions and the estimated volume.

The regularization term

$$\Psi = \parallel \nabla I_f \parallel^2 \tag{3.8}$$

is defined to ensure both the solution and the convergence to a global minimum. A weighting factor λ is defined to balance the contribution of the terms Q and Ψ, defining the global energy E as:

$$E = Q + \lambda \cdot \Psi \tag{3.9}$$

The optimal estimated volume I_f is the one that has the minimal energy E, $I_f = \text{argmin}_{I_f} E$. It can be expressed as:

$$I_f = argmin \left(\sum_i \parallel I'_i - h'_i * I_f \parallel^2 + \lambda \cdot \parallel \nabla I_f \parallel^2 \right) \tag{3.10}$$

where as mentioned in Soler et al. (2005b), the authors have chosen to initialize the system with the average of the all the input images.

Soler et al. (2005a) has shown this technique has comparable performance to Generalized Averaging. It was only been applied to two image with a very close view angle. However, this algorithm can be used for more than two images.

Wavelet method

The Wavelet method (Rajpoot et al., 2009; Szmigielski et al., 2010) uses a wavelet transform to separate the image information into low and high-frequency sub-bands, and aims to deal with the image feature and noise pattern separately.

Firstly, the technique applies forward wavelet transform W^F on image I'_i to separate the image information into low and high frequency signals

$$W^F(I'_i) = [Wav^i_L, Wav^i_H] \tag{3.11}$$

The assumption is that the low frequency wavelet signal (Equation 3.12) contains information of anatomical features, which need to be kept and

passed to final compounded image, and the high frequency wavelet signal (Equation 3.13) contains information of image noise, which needs to be removed in the final compounded image. Therefore, the Maximum operator is used on the low frequency wavelet signals and Mean operator is used on these high frequency wavelet signals.

$$Wav_L^f = max\{Wav_L^1, ..., Wav_L^N\} \tag{3.12}$$

$$Wav_H^f = mean\{Wav_H^1, ..., Wav_H^N\} \tag{3.13}$$

Finally, by applying the inverse wavelet transform W^I to the processed low and high frequency signals (as calculated in equation 3.12 and 3.13), the final compounded image is re-constructed

$$I_f = W^I[Wav_L^f, Wav_H^f] \tag{3.14}$$

The Wavelet based compounding method separates each of the aligned images into low and high-frequency components (anatomical feature and noise patten accordingly), thus allowing to use relevant fusion rules. This algorithm can be applied to any number of images.

3.2.3 Ultrasound specific methods

As described in Chapter 2, echo imaging has its own characteristics due to ultrasound physics and the way ultrasound interacts with tissue. For example, reflection is a function of the interface acoustic impedances, and the strongest reflection is expected when the boundary surface is normal to the ultrasound beam direction; homogenous tissue regions contain speckle patterns caused by constructive and destructive interference between waves

reflected from the numerous small scale structures within homogeneous tissue; artefacts commonly exist in echo images, and will cause missing anatomical features or false information being introduced. Therefore, a number of compounding algorithms have been specifically proposed for ultrasound.

Incidence Angle Weighting

As ultrasound wave reflection is angular dependent, Incidence Angle Weighting (Leotta and Martin, 1999) weights image information based on the angle between the ultrasound beam and the reflecting surface normal.

The authors proposed that, according to the viewing geometry, the echo image is expected to have the highest intensity value at the incident angle close to zero degrees, i.e. approximately normal to the surface. Therefore, it could be advantageous if images with different incident angles contribute differently to the final volume, where each intensity value is weighted by the calculated incident angle of ultrasound beam.

$$I_f = \frac{\sum_{i=1}^{N} w_i(\alpha) * I'_i}{\sum_{i=1}^{N} w_i(\alpha)} \tag{3.15}$$

where α is the incident angle and $w_i(\alpha) = \cos(\alpha)$ is the weighting for the i^{th} intensity value.

The authors showed that image quality has been enhanced using incident angle weighting compounding, compared to the maximum and mean methods using phantom data (simple bone structures). This method was used to compound three images and the algorithm has no limitation to be applied to larger numbers of images. However, it is based on obtaining the information on surface normal direction (by deriving surface orientation from the image gradient), and the quality improvements vary depending on

the particular target. Obtaining the surface orientation from highly reflecting static surfaces (bone phantom) is easier than from low reflecting moving surfaces. Therefore, it could be difficult to apply the technique to complex objects, such as cardiac images, where complex structures are presented and segmentation is a challenge.

Phase-based method

The Phase-based method (Grau and Noble, 2005) uses image phase information to calculate feature importance and local orientation parameters which are used as weights for each intensity value in the image set. As an ultrasound specific method, it works on the following principles: In regions which contain homogeneous tissue, the intensities from multiple images are averaged to improve SNR, and at strongly reflecting interfaces the input from different images is weighted based on the angle between the ultrasound beam and the reflecting surface normal. Therefore, the method needs to determine when different compounding strategies should be used, and at interfaces the method needs to calculate surface normal information.

The authors calculate phase using the monogenic signal. A detailed explanation of monogenic signal can be found in Felsberg and Sommer (2001). Then the following steps are carried out:

First of all, phase congruency (Morrone and Owens, 1987) is calculated.

$$PC = \max_{\varphi \in [0,2\pi]} \frac{\sum_s A_s \cos(\varphi_s - \bar{\varphi})}{\sum_s A_s} \tag{3.16}$$

where A_s represents the amplitude and φ_s is the local image orientation of the s^{th} Fourier component of signal.

In order to have the significant information transferred into the final

volume, weightings S_i are used to determine how important the information is in the image i at location $\mathbf{x}$.

$$S_i = f(PC_i) \tag{3.17}$$

where any monotonic function f can used with $f(0) = 0$ and $f(1) = 1$.

If S_i is high in more than one image, then S_i is not sufficient, so another measure M_i is introduced.

$$M_i = \cos(\varphi_i - \theta) \tag{3.18}$$

where θ is the incident angle of ultrasound beam.

In the case where three images are combined, the weight w_i for image i can be calculated as:

$$\begin{aligned} w_i &= [S_1\bar{S}_2\bar{S}_3 + S_1S_2\bar{S}_3M_1\bar{M}_2 + S_1\bar{S}_2S_3M_1\bar{M}_3 + S_1S_2S_3M_1\bar{M}_2\bar{M}_3] \\ &+ \frac{1}{2}[S_1S_2\bar{S}_3M_1M_2 + S_1\bar{S}_2S_3M_1M_3 + S_1S_2S_3M_1(M_2\bar{M}_3 + \bar{M}_2M_3)] \\ &+ \frac{1}{3}[S_1S_2S_3(M_1M_2M_3 + \bar{M}_1\bar{M}_2\bar{M}_3)] + \frac{\rho}{3}[\bar{S}_1\bar{S}_2\bar{S}_3] \end{aligned} \tag{3.19}$$

A coefficient ρ has been introduced here for noise reduction. In the algorithm, the weighting factors are calculated at each voxel position for each image i. Finally, the compound image is derived by weighted mean (equation 3.15).

This technique was the state-of-the-art 3D echo compounding method when I started my PhD research. However, the algorithm has only been designed for a maximum of three volumes. Due to the complicated calculation of the weighting measure (equation 3.19), this method is difficult to be extended for larger numbers of images.

3.2.4 Summary

In this section, previous 3D compounding methods have been reviewed using basic statistical, advanced and ultrasound based strategies. An overview of these prior work is also shown in Table 3.1. They have typically been applied to a relatively small number of images (most of them use two images, one uses six). All the methods report improvements in image quality i.e. improving SNR and contrast, and extending the field-of-view. Although previous compounding work hasn't directly claimed or demonstrated a reduction in the effect of echo artefacts, fundamentally, some of the algorithms published in the literature can handle artefacts. For example, the maximum method will always pass the highest intensity value to the compounded image, so that the low intensity shadowing artefacts will be reduced; the generalized averaging method will use the maximum operator to keep a feature if it is present in only one image, so again the shadowing artefacts will be reduced in this situation; The phase-based method has the mechanism of detecting important anatomical features where shadowing artefacts will be given a low weighting value throughout the compounding process.

3.3 Image registration

Compounding requires the multiple views of same anatomical feature to be aligned. Normally, echocardiographers acquire data from standard view angles, so corresponding anatomical features usually present in the echo images at different positions and orientations. Image registration is used to spatially align images, so that corresponding features can be related. In the

Table 3.1: An overview of the prior 3D echo compounding work.

	Leotta and Martin 1999	Soler *et al.* 2005	Grau and Noble 2005	Szmigielski *et al.* 2010
Compounding method	Incident angle weighing	Generalized averaging; Multiview deconvolution	Phase-based compounding	Wavelet
Registration method	Magnetic tracking	Manually defined initial estimation followed by image based registration with a gradient decent algorithm	Landmark based initial alignment followed by a manual fine correction	Image based registration with normalized cross-correlation as similarity measure
Intensity normalisation	No	No	No	No
Directly addressing artefacts	No	No	No	No
Simulated data	0	0	1	0
Phantom data	2	0	0	2
Volunteer data	2	0	2	16
Patient data	0	8	0	16
The maximum number of images being compounded	3	2	2	6
Numerical assessment	peak grey level; border width	contrast; SNR	contrast to noise ratio	contrast to noise; SNR
Visual assessment	No	No	No	Yes

literature, good review articles are available and a number of image registration methods are described: for general purpose registration, see Zitova and Flusser (2003), for medical image registration see Maintz and Viergever (1998) and Hill et al. (2001), and for methods focused on cardiac images see Makela et al. (2002). In this section, general and ultrasound specific registration strategies are introduced and methods for multiple image alignment are discussed.

3.3.1 Components of a general image registration algorithm

Various medical image registration methods have been reviewed by van den Elsen et al. (1993), Maurer et al. (1993) and Maintz and Viergever (1998). Methods can be classified into 3 types: extrinsic, intrinsic, and non-image based. Extrinsic methods requires foreign objects to be included in the image (markers attached to the patient). Intrinsic methods are based on the image content only, which can be feature based (a set of identified landmarks, segmented structures, or surfaces, etc.) or directly based on the measurement from the image intensity values (voxel-based). Non-image based methods can be used to carry out multimodal image registration. These methods typically use a tracking system and a calibration process which determines the relationship between the image coordinate systems and this tracking system. Examples are PET-CT scanners and the XMR system which is used to align fluoroscopy with MR image (Rhode et al., 2003).

A typical registration algorithm consists of the following components:

Table 3.2: Overview of transformations and their degrees of freedom (DOF)

Transformation	DOF
Rigid	6
Translation	3
Similarity	7 or 9
Affine	12

Type of transformation : This refers to the transformations allowed by the registration algorithm to align the images. If the images registered are of the same object that is merely in a different position (i.e. no deformation has occurred), then the transformation can be described using just translations and rotations. This gives us a rigid-body transformation which although the heart is obviously a non-rigid structure is the transformation commonly used to align images acquired at the same cardiac phase. A quick overview of transformations and their degrees of freedom (DOF) can be found in Table 3.2.

Similarity measure : This provides a numerical value indicating, for a given transformation between the images, how well aligned the images are.

Optimization : To achieve registration the algorithm needs to alter the degrees of freedom which parameterise the transformation in order to achieve an optimum value for the similarity measure.

Maintz and Viergever (1998); Hill et al. (2001) and references therein have proposed several voxel-based similarity measures. Authors stated that corresponding voxels in two images are related by a certain measure, or statistical relationship. Taking two images for example (pair-wise registra-

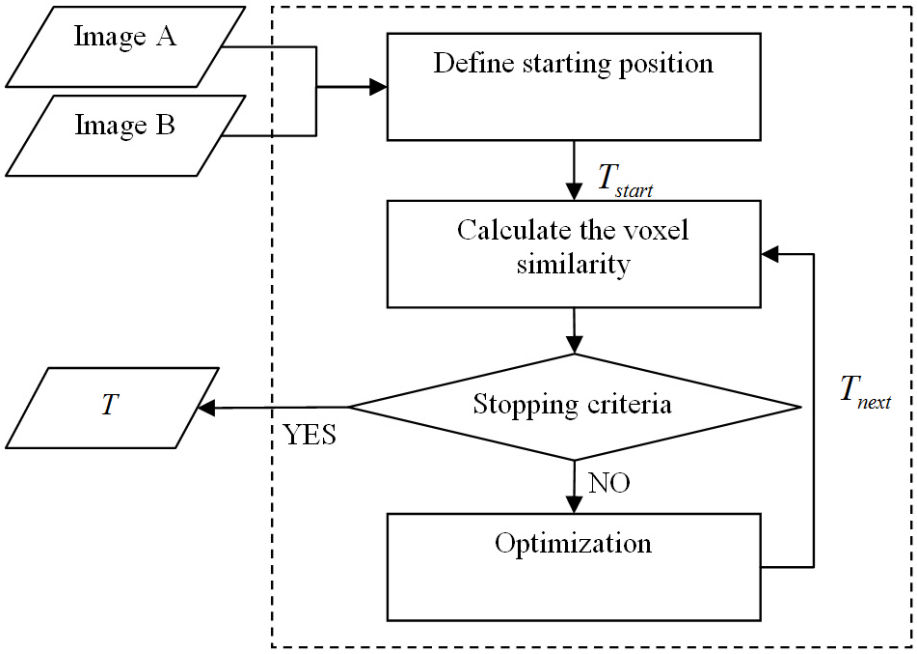


Figure 3.2: Flowchart to demonstrate how the components in a typical voxel-based registration algorithm are used to register two images (A and B). T is the transformation output between image A and B. T_{start} is the starting position. T_{next} is the next iterative transformation carried out by optimization procedure.

tion), the work flow of a typical voxel-based registration algorithm has been shown in Figure 3.2.

3.3.2 Image registration strategies used by prior compounding methods

Image registration algorithms have been applied to match ultrasound images together for a number of applications. Research has been carried out to produce extended field-of-view images (Poon and Rohling, 2006; Wachinger, 2007); removing deformation artefacts in 3D freehand scanning (Treece et al., 2002) and to facilitate sensorless freehand tracking (Gee et al., 2006); as a precursor to compounding (Grau et al., 2007); and to detect intraoperative deformations.

Compared to other imaging modalities, ultrasound has unique characteristics in terms of noise, angular dependent reflection, and artefacts, which makes it particularly difficult to use voxel-based similarity measures. As stated in Makela et al. (2002), not many publications exist on cardiac ultrasound image registration, due to the quality of the images. In Shekhar et al. (2004), mutual information was used to register rest and peak stress echo images. In Ledesma-Carbayo et al. (2005), mean square difference was used as a similarity measure for registration of echo images for motion tracking. In Wachinger et al. (2012), nakagami-based similarity function was proposed for ultrasound image alignment.

In this section, the following image registration strategies are discussed from the 3D echo compounding literature:

Non-image based : Leotta and Martin (1999) use magnetic tracking information to align the images. The magnetic tracking system functions by measuring the strength of the magnetic fields generated by a transmitter. It doesn't require a line of sight between transducer and receiver. Therefore, it allows the transducer have a wide range of orientations. Additionally, the receiver coils can be placed close to the transducer face, which should increase accuracy. However, the accuracy of magnetic tracking system can be affected by the variance of the external magnetic field, particularly caused by the presence of metallic objects. Furthermore, patient movement or breathing motion will affect the accuracy of the result if only tracking information is used to align images, and the registration errors may cause problems in further processing.

Intrinsic - feature based : Grau and Noble (2005) use manual registration based on manually selected landmarks. This obviously involves human intervention which can be time consuming. Also Soler et al. (2005a) use feature-based surface matching to manually define their initial estimation for registration algorithm. Although the process is adequate to show the effectiveness of subsequent compounding, for a system in routine clinical use a more automated registration algorithm would probably be necessary.

Intrinsic - voxel based : Soler et al. (2005a) use a gradient decent algorithm to find a rigid transformation. The algorithm minimizes an intensity-based distance metric, where a robust estimator is applied to the subtraction of the corresponding intensities between two images. Also Rajpoot et al. (2009); Szmigielski et al. (2010) use normalized cross-correlation as similarity measure to align images with close view angles. The rigid transformations have been used and the algorithm is automatic.

3.3.3 Registration for multiple images

In the situation where multiple images need to be aligned, the optimisation strategies can become more complex and very computationally expensive.

In literature, registration for multiple images is mostly used for generating an atlas using a set of MR images. The atlas can be created using a number of methods, probably the most basic one is by registering all the images to a chosen reference and averaging the resulting images; more advanced method is to use transformation averaging to generate an average

atlas space (Aljabar, 2008).

For aligning multiple ultrasound images, one solution is to use multiple implementations of pair-wise registration (Figure 3.3 top left) where all the images are registered to the reference image independently. However, the result is highly depending on the choice of reference image, i.e. the quality and field-of-view of the reference image can greatly affect the registration result. There is another solution, group-wise optimisation strategies which are proposed in literature to measure the similarity of a group of images. The simultaneous optimisation (Wachinger and Navab, 2009; Wachinger et al., 2007) involves all the images rather than just on two, and directly optimizes the global transformations for all the images simultaneously (Figure 3.3 top right). The result is independent of the choice of the target image. Compared to the multiple implementation of pair-wise registration, the group-wise optimization is expected to produce more robust and accurate results for the compounding process.

There are two approaches for the simultaneous optimisation which have been proposed by Wachinger (2007), which vary in their computational complexity. The full-simultaneous strategy considers all available images simultaneously during the optimisation process. It optimizes over $6 \cdot N$ parameters in each step for all the N available images at the same time which can be very computationally expensive. The semi-simultaneous strategy (as shown in Figure 3.3 bottom) is the modification which just optimizes 6 parameters of a single image at a time until convergence, the method then goes to each following individual image in turn until the positions of all the images have been optimised for a cycle over all images. If the

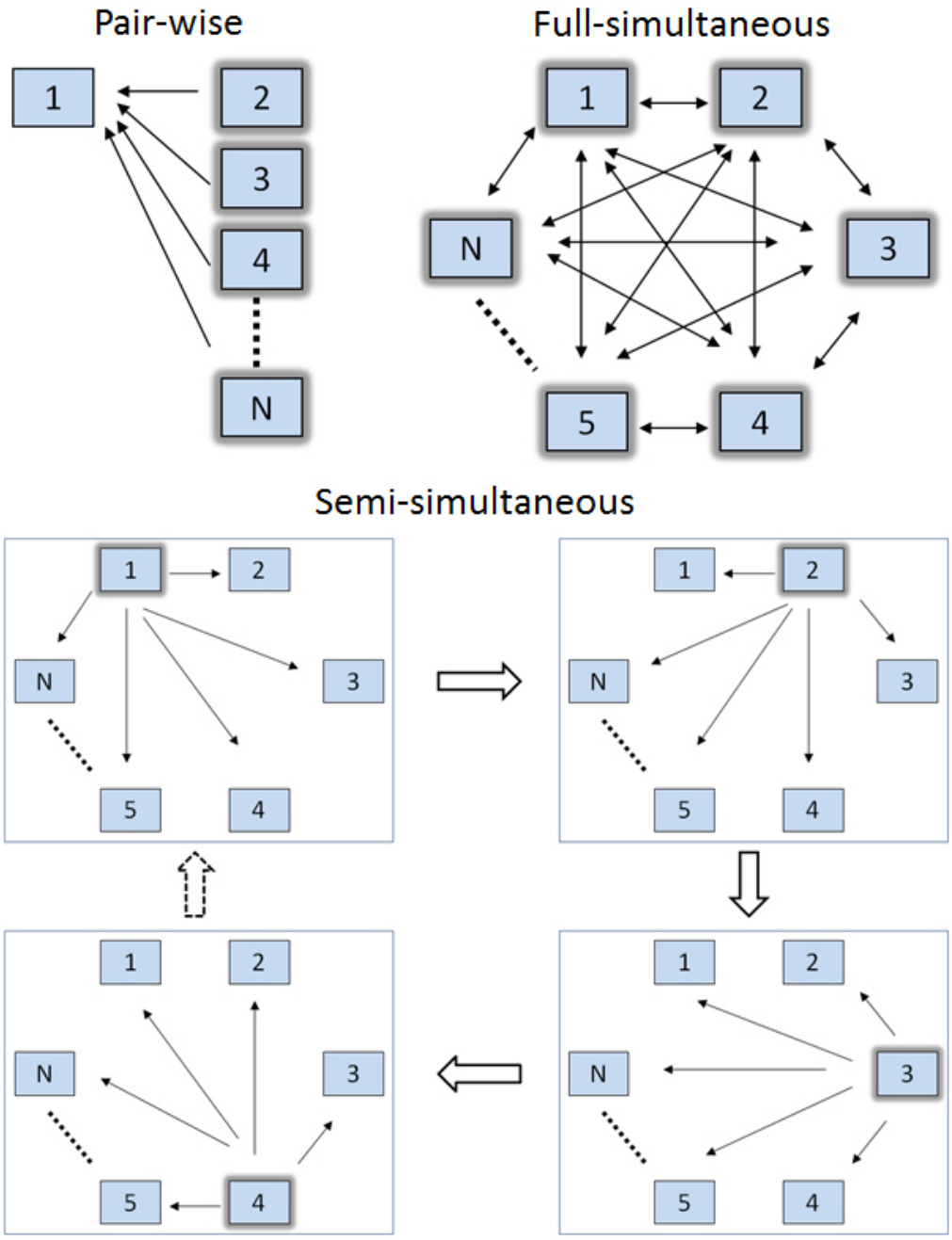


Figure 3.3: Illustration of pair-wise and group-wise implementations used to register multiple images. In the situation of aligning N images, the boxes with highlighted edges denote the moving images during the optimisation process, and the arrows show the transformations calculated between images. (top left) Multiple implementations of pair-wise registration is highly depending on the choice of reference image (Image 1 in this case). (top right) Full-simultaneous group-wise strategy optimizes all the images simultaneously. (bottom) Semi-simultaneous group-wise strategy takes each image in turn, and optimizes the transformation of that moving image to all the other images in each step.

global alignment hasn't been achieved, a new cycle is started. Otherwise the algorithm terminates. It has shown to be a good trade off between registration accuracy and computation time.

3.4 Validation strategies

A number of different validation methods have been proposed to evaluate the performance of prior compounding strategies, with respect to the numerical measurements, the visual inspection and the testing datasets.

In terms of the numerical measurements, quantitative parameters commonly used to determine the improvement of image quality in the literature are field-of-view extension ($\triangle FOV$), contrast increase ($\triangle Contrast$), signal-to-noise ratio improvement ($\triangle SNR$) and border measurement such as peak grey level and full-width at half-maximum (FWHM). Most of the previous compounding work (Leotta and Martin, 1999; Soler et al., 2005a; Grau and Noble, 2005; Rajpoot et al., 2009; Szmigielski et al., 2010) have used the signal-to-noise ratio and contrast as the numerical measurements. These are useful as they are standard measurements, and can allow some comparison between algorithms in the literature.

Quantitative measurements only show particular aspects of image quality and clinicians are often used to judge the quality by visual inspection (qualitative measurements). Whatmough et al. (2007) asked observers to review 2D compounded ultrasound images. High quality images were identified and the effect of different compounding settings were evaluated. Observers perceived that compounded images had improved contrast resolu-

tion, border definition and SNR. However, it could be difficult for observers to distinguish the differences between spatial resolution and image quality, where they may have misinterpreted a reduction in the speckle of a compound image as motion blur.

Both numerical measurements and visual inspection can be used to assess the global and local image quality, but depend on the application. For example, if artefact regions or specific anatomical features are considered, a local region measure is used (Soler et al., 2005a); If an overall image quality or field-of-view extension need to be defined, the global information is used (Soler et al., 2005a; Rajpoot et al., 2009).

Various testing datasets have been used in the previous compounding work, to verify the effect of ultrasound image quality improvement:

Simulated data : For the convenience of preliminary evaluation and in order to have more control of generated data (e.g. generate speckle noise), Grau and Noble (2005) used Field II (Jensen, 1996) to produce synthetic data to test the basic functionality of their compounding method. However, simulated images can be open to criticism due to the realism of the simulation.

Phantom data : Phantom data (Leotta and Martin, 1999; Szmigielski et al., 2010) is a compromise between the control over imaging and obtaining the real images from a ultrasound machine. However, phantom can be too simplistic in terms of the complexity of structures and possibly lack of motion.

Volunteer data : In order to reflect the realistic scenarios, real echo im-

ages of volunteers are used to validate different compounding methods (Grau and Noble, 2005; Rajpoot et al., 2009; Szmigielski et al., 2010). It is not too difficult to acquire as many images as possible from volunteers, and there are also no time restrictions.

Patient data : Soler et al. (2005a); Szmigielski et al. (2010) used patient data for validation. Acquiring patient data is similar to volunteer data. However patient cases are more complex than healthy volunteers. For example often the FOV is more restricted, the images acquired can be of lower quality, and time and resource restrictions faced by the clinics and their clinicians.

3.5 Conclusion

In this chapter, I have reviewed previous compounding techniques and discussed compounding algorithms, image registration methods and validation strategies accordingly.

The prior compounding methods have shown the ability to remove ultrasound noise, and enhance the boundary definitions. However, artefacts are common in echo images and can have a huge effect on image quality, such as missing useful anatomical structure or false information generated. None of the previous methods has directly addressed the effect of artefacts. Also previous algorithms have only been applied to no more than six images.

In the next chapter, I investigate the use of compounding using more images. Both statistical and modified state-of-the-art compounding methods are implemented.

Chapter 4

3D echo compounding with larger numbers of images

This chapter presents my study on compounding larger numbers of 3D echo images. It starts by discussing the possible advantages and disadvantages of using more images. Then my proposed methodologies to be applied to larger numbers of images are described. The details of experiments and results will be presented, and benefit and potential issues of using more images will be discussed in the end.

4.1 Motivation

There are potential advantages of using more 3D echo images for compounding. First of all, there is evidence of improvement in SNR when compounding larger numbers of 2D images from a static subject (Rohling et al., 1997). The improvement has been demonstrated as being proportional to $\sqrt{N}$ for N uncorrelated 2D images. Secondly, some parts of the heart are more

visible from particular view angles, which is due to the angular dependent effect of ultrasound imaging. Multiple images covering a wide range of views will have the ability to fill in the missing image information and are better able to counter the angular dependent effect. In addition, as it is relatively easy to acquire more images with an echo machine, so it should possible to benefit from using all the data available, without substantially altering clinical practice.

However, there are several potential disadvantages of using more images. Firstly, registration errors between volumes could be accumulated when more images are compounded. For example, when using the maximum method to compound images containing registration errors, the anatomical feature in the output image might become thicker; or when using the mean compounding method with registration error, the output feature will be blurred due to the intensity averaging on the displacement feature from different images. Therefore, the compounded image quality could be affected by using more images if registration errors occur. Secondly, it may not be feasible in clinical situations to acquire a larger numbers of images from different view positions, due to patients having limited or poor acoustic windows. Thirdly, it would take more time from data acquisition to finally displaying images, due to the time required for acquiring more images, and longer image registration and compounding processes.

In the following sections, the effect of using more 3D echo volumes on image quality will be investigated.

4.2 Methodologies

In this section, methods will be proposed to compound larger numbers of 3D echo images. The phase-based compounding method (Grau and Noble, 2005) was the state-of-the-art compounding technique of at the beginning of my PhD. However, as Section 3.2.3 described, the original implementation was for use with up to 3 images, and the extension of phase-based compounding to larger numbers of images is non-trivial due to the process required to calculate the weights. In this section, the phase-based compounding method is extended to use more images. It is extended using two different approaches, and compared with the two basic mean and maximum statistical methods.

4.2.1 Extension of phase-based method

The following describes the two different ways I proposed to allow the phase-based approach to be easily applied to large sets of images:

Phase Grouped compounding : The images are divided into groups of 3 in the order of data acquisition and each group is compound using the standard phase-based method (Figure 4.1). This procedure is then repeated until a single compounded image is produced.

Phase Selection compounding : This method uses an image selection step (Figure 4.2) as a plug-in to enable the extension of standard phase-based method. This has some analogies with the "Lucky Imaging" techniques, which have been used in astronomical photography (Law, 2006). The quality of telescope images change rapidly due to the

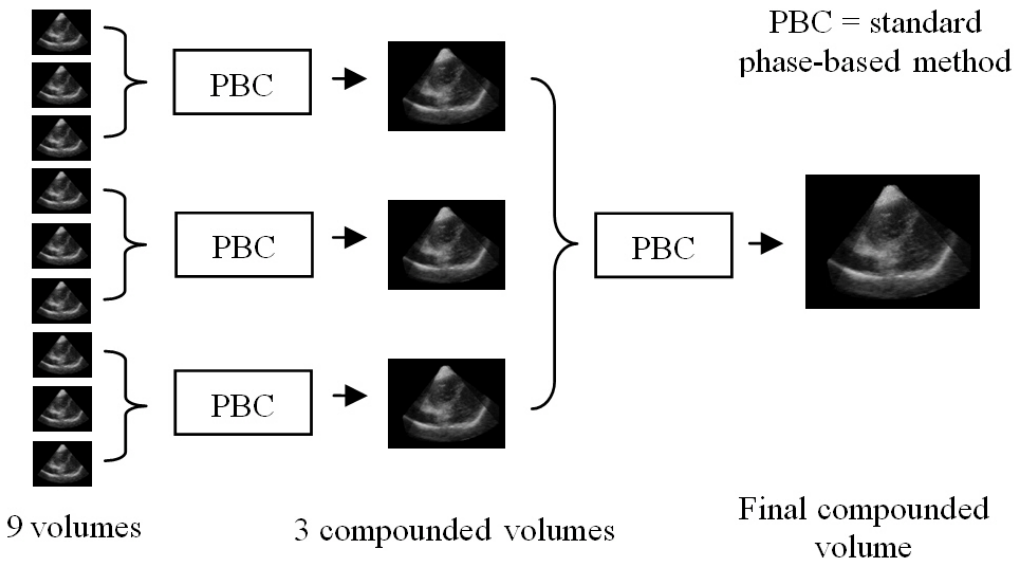


Figure 4.1: Phase Grouped: my proposed extension to standard phase-based compounding methodologies is illustrated using 9 images. Images are repeatedly grouped into sets of 3 images until it ends up with a single image.

variance of the density fluctuations in the atmosphere. Lucky imaging selects and combines images which are less affected by the atmosphere. In my case, a selection strategy is used to choose a subset of values which are both "representative" intensities and also unlikely to contain artefacts. These selected intensities are then used as input to the standard phase-based method. In this chapter, a sample selection strategy is used, which chooses the three highest intensity values for each voxel location. This is because many image artefacts result in shadowing, and so create low intensity values. Therefore, my assumption is that selection of high intensity values should act as a simple artefact removal filter.

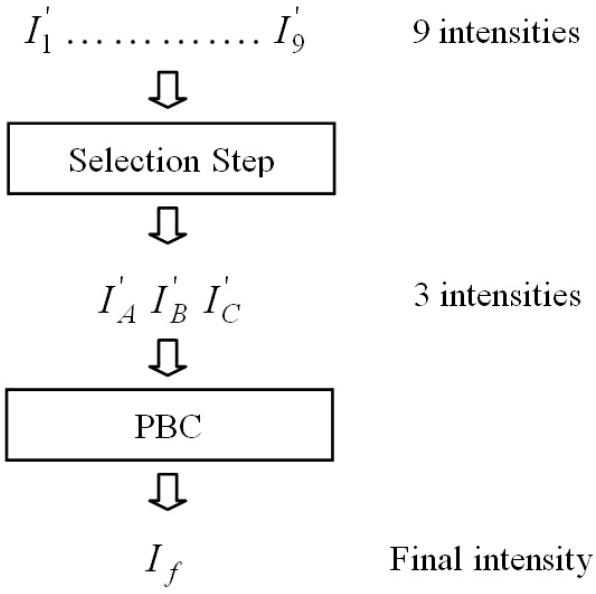
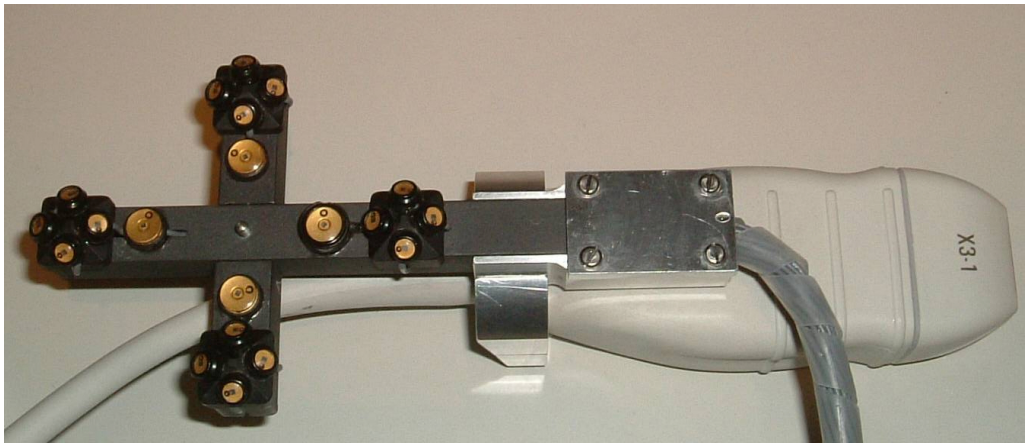


Figure 4.2: Phase Selection: my proposed extension to standard phase-based compounding methodologies is illustrated using 9 images. It selects the 3 most "representative" images for each voxel in I_f .

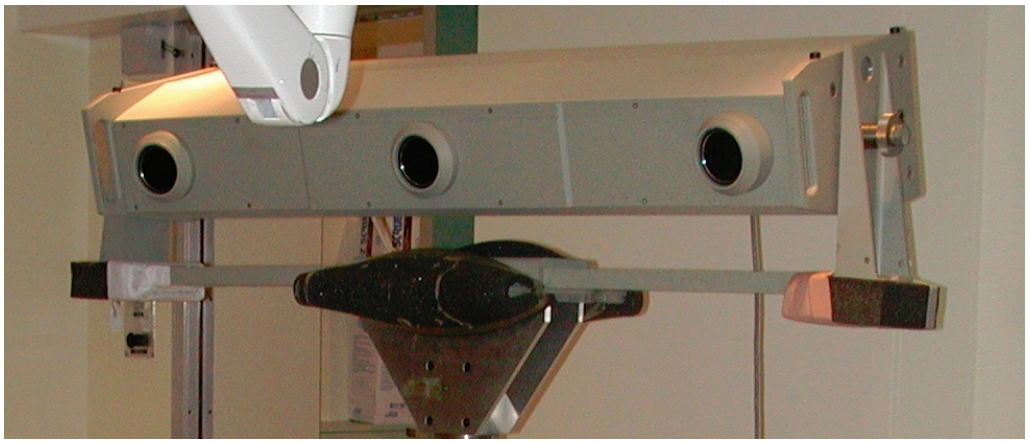
4.3 Image registration

In order to carry out image compounding, the transformation between each of the images and a reference coordinate system must be accurately calculated. Registration errors will cause problems for compounding such as anatomical feature blurring. For the experiments described here, spatial compounding of 3D echo images is my main focus. The ECG is used for temporal gating. Then the images at the end-diastolic phase are chosen because the heart is at the most relaxed stage and the effect of cardiac motion will be minimized for the image registration and compounding.

My registration method uses an initial starting position obtained by optically tracking the transducer. As shown in Figure 4.3 an array of light emitting diodes is attached to the 3D X3-1 matrix array echo transducer, which allows the transducer to be tracked using a Northen Digital Optotrak (NDI, Ontario, Canada).



(a)



(b)

Figure 4.3: (a) an array of light emitting diodes is attached to the 3D X3-1 matrix array echo transducer; (b) Northen Digital Optotrak is used to capture the movement of the transducer.

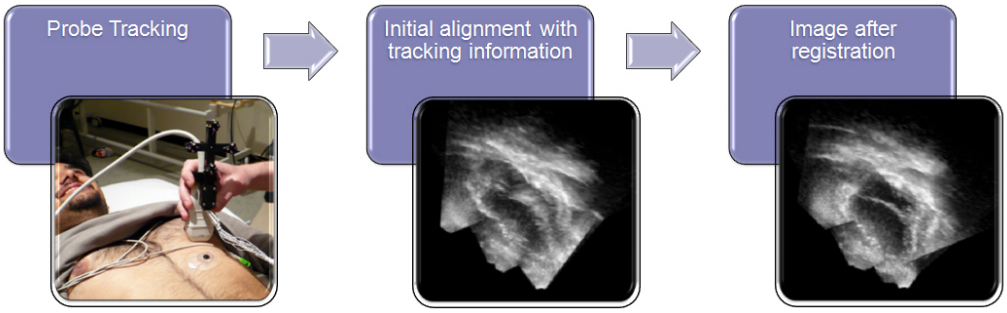


Figure 4.4: Overview of image registration procedure. Images are acquired with a tracked and calibrated echo probe (left); They are aligned using the tracking information only (centre) and feature blurring can be seen due to respiratory motion or patient movement. The initial alignments can be used as a starting estimate for image-based registration to improve compounded image quality (right).

Although optical tracking will be able to accurately localise the transducer, the heart will move due to respiratory motion or patient movement. Use of optical tracking alone would result in large registration errors (Figure 4.4). Therefore, I use the optical tracking as an automated method to provide a starting estimate for an image-based registration.

My image-based registration algorithm uses an independent implementation of the phase-based similarity measure (Grau et al., 2007) with the semi-simultaneous strategy as discussed in Section 3.3.3 to align larger numbers of images. The phase-based similarity is based on the local orientation and phase information extracted from images to guide the registration. This similarity measure is specially designed for ultrasound images as phase information is unchanging to the variance in brightness and contrast of the image.

In order to compute the local orientation and (local) phase of the 3D data I use the monogenic signal analysis, as proposed in Felsberg and Sommer (2001). The value of the monogenic signal of an three dimensional signal

$I(\mathbf{x})$, at location $\mathbf{x}$, consists of four elements. The first element is the value of the original signal and the remaining three elements are the Riesz transform of the signal. In the case of 3D data the Riesz transform can be calculated by filtering the signal with three oriented filters with a frequency response described by H_1, H_2 and H_3.

$$H_1(u, v, w) = \frac{u}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \tag{4.1}$$

$$H_2(u, v, w) = \frac{v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \tag{4.2}$$

$$H_3(u, v, w) = \frac{w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \tag{4.3}$$

where u, v and w are the coordinates in Fourier domain. In practice I first filter the signal with a bandpass filter so I can work on the scale which is most meaningful for the data. As in Grau et al. (2007) a log Gabor filter was used. The ratio of the standard deviation of the Gaussian, which is describing the log Gabor filter's transfer function in the frequency domain to the filter center frequency, was chosen equal to 0.65.

The local orientation and phase of the signal in spatial location $\mathbf{x}$ is computed from the four dimensional value of the monogenic signal in this location. The local orientation at location $\mathbf{x}$ in scale c is calculated according to:

$$o(I, c) = \frac{(g_c * h_1 * I, g_c * h_2 * I, g_c * h_3 * I)}{(\sum_j (g_c * h_j * I)^2)^{\frac{1}{2}}} \tag{4.4}$$

where I is the original image signal, g_c is the impulse response of the band-pass filter for scale c and h_1, h_2 and h_3 are the impulse responses of the three oriented filters used for the computation of the Riesz transform. The

local phase at location $\mathbf{x}$ is computed according to:

$$\varphi(I,c) = atan2\left((\sum_j (g_c * h_j * I)^2)^{\frac{1}{2}}, g_c * I\right) \tag{4.5}$$

Since all the 3D data in the study were of the same resolution, a single scale approach was used. I chose to compute the local orientation and phase in scale level 3.

The phase-based registration implemented by Grau et al. (2007) was used to align a pair of images from each of their seven datasets. These aligned images were then presented to an echo expert for visual assessment of registration accuracy. The heart was split up into segments using the standard 16-segment model. For each segment the expert measured the misalignment visually and scored them on a scale from worst (error $> 5\,mm$) to best (error $< 2\,mm$). In total 85% of the segments had the registration error less $2\,mm$.

The semi-simultaneous strategy works as follows: The method begins with all the images in an initial rough alignment achieved by optical tracking. Then each image was taken in turn and to calculates the value of the phase-based similarity measure (an average value of the similarity measures between this image and all the other images), and each of the six rigid body parameters altered by a stepsize of $2\,mm$ or $2\,^\circ$. If an improvement in the similarity measure was found, then the image was moved. If no improvement is found for any of the images then the stepsize was reduced by a factor of two until a stepsize of $0.0625\,mm$ or $0.0625\,^\circ$ is reached when the algorithm stops.

Prior to registration, the echo volumes were pre-processed as follows: An anisotropic diffusion filter (Perona and Malik, 1990) was used to reduce

image noise while maintaining feature boundaries. The images were reduced in resolution to 2 mm cubed voxels by blurring and subsampling in order to decrease computation time. These image processing steps were only applied for the registration algorithm, and not for the compounding experiments. The alignment process was totally automatic.

4.4 Experiments

The experiments investigate the effect of compounding increased numbers of images (i.e. 1 to 10) within each dataset, using different compounding methods: Mean, Maximum, Phase Grouped and Phase Selection. The evaluations have also been carried out on the consistency and capture range of the semi-simultaneous registration proposed in the previous section.

4.4.1 Data acquisition

In general, to obtain clinically useful echo images the clinician will manoeuvre the probe to obtain good acoustic windows. A number of acoustic windows are used: parasternal, apical and subcostal views. This process also usually involves the patient changing position and holding their breath at different respiratory phases. In addition there are a number of imaging parameters that can be optimised on the echo machine, such as depth, overall brightness and time-gain compensation. Throughout the imaging process ECG gating is used to label the corresponding cardiac phases between acquisitions. Echocardiographers have a great deal of skill and experience in altering all these factors to achieve the best possible image quality from

the patient. I have attempted to design my compounding system so that I can take advantage of the echocardiographers skill as much as possible, however, I have found it necessary to impose some restrictions to acquire images suitable for my compounding algorithm. Firstly I requested that the only changes to machine settings between image acquisitions were either depth and/or overall brightness (gain). Secondly I requested that the volunteer was asked to move as little as possible during image acquisition. This was necessary as my registration algorithm (described in Section 4.3) requires a starting position, which was achieved for most of my experiments by optically tracking the transducer. When movement was necessary the image number was recorded and a manual alignment method was used to update subsequent starting positions.

In this chapter, three sets of 3D echo volumes (which will be referred to as Dataset-Q1, Q2, Q3) were acquired by experienced echocardiographers from volunteers using a Philips iE33 ultrasound system with a 3D X3-1 matrix array transducer. The transducer was optically tracked as described in Section 4.3. Probe calibration was performed using the method described by Ma et al. (2008). Triggered by ECG gating, wide sector acquisitions (during 4 or 7 heartbeats) were taken with volunteers positioned in the left lateral decubitus position. The angle and position of the probe was slightly changed (around $1-2\,cm$) for each acquisition to vary the content of each volume. Both parasternal and apical views (in total 10 images for each dataset) were acquired. The reason of acquiring only 10 volumes for each dataset is that I estimate that 10 is near to the upper limit of acceptance and could be obtained in approximately 8 minutes by an experienced clinician

following a devised protocol in a routine clinical practice. Information on the quality of individual images and whether or not images were acquired from challenging acoustic windows for each healthy volunteer were not recorded.

4.4.2 Experiment to determine semi-simultaneous registration consistency and capture range

In order to compound multiple images, the corresponding anatomical feature needs to be aligned to the same coordinate system. The accuracy of the mapping is important and registration errors will likely lead to errors in the final compounded images. It is therefore important to know how accurate my registration method is able to align the set of input images. In addition to registration accuracy my system needs to work robustly and have a large enough capture range to enable automatic registration from a suitable starting position (i.e. using information from an external tracking device). This section describes experiments to calculate registration consistency and capture range for the semi-simultaneous registration method used in this Chapter.

Method for consistency experiment

Ideally I would like to know the accuracy of my registration method when applied to clinical data. However, accuracy measurements typically requires either a ground truth registration or landmarks and these are very difficult to obtain accurately in echo images of patients or volunteers (Grau et al., 2007; Zhang et al., 2007; Guo et al., 2009; Dalvi et al., 2010; Karnik et al., 2010). Simulated data contains an inherent gold standard, however it is

very difficult to simulate realistic echo images which include all the aspects of echo image quality which could affect registration. For these reasons I propose to calculate registration consistency using real echo images rather than a measure of accuracy, such as the Target Registration Error (TRE).

My consistency method is based on the method described by Holden et al. (2000). In Holden et al. (2000), this method (illustrated in Figure 4.5) is used to calculate the consistency between three images (image 1, 2, 3) which have been registered together to calculate three independent transformations (T_{1-2}, T_{2-3}, T_{3-1}), between each image pair such that T_{1-2} transforms image 1 into the coordinate system of image 2. The consistency measure works by transforming a point (or set of points) around a closed loop in the set of images. As shown in Figure 4.5, the position p is transformed around the closed loop to $p' = T_{3-1} \cdot T_{2-3} \cdot T_{1-2} \cdot p$. Then if all the registrations are perfectly accurate $p' = p$. However, registration errors will result in there being a difference in position p' and p, and the consistency value calculated is the difference in these positions $E = \|p' - p\|$, or if set of points is used, a mean displacement error $MDE = \frac{1}{M} \sum_{j=1}^{M} E_j$ is calculated for these M number of points.

The method described in Holden et al. (2000) was only used when registrations were between pairs of images. In my registration (semi-simultaneous registration) the algorithm registers multiple images at the same time. The semi-simultaneous registration system is designed so that within a set of multiple images all of the individual transformations are automatically consistent. In order to keep the transformations independent in the consistency check, a revised method was carried out which is illustrated in Figure 4.6.

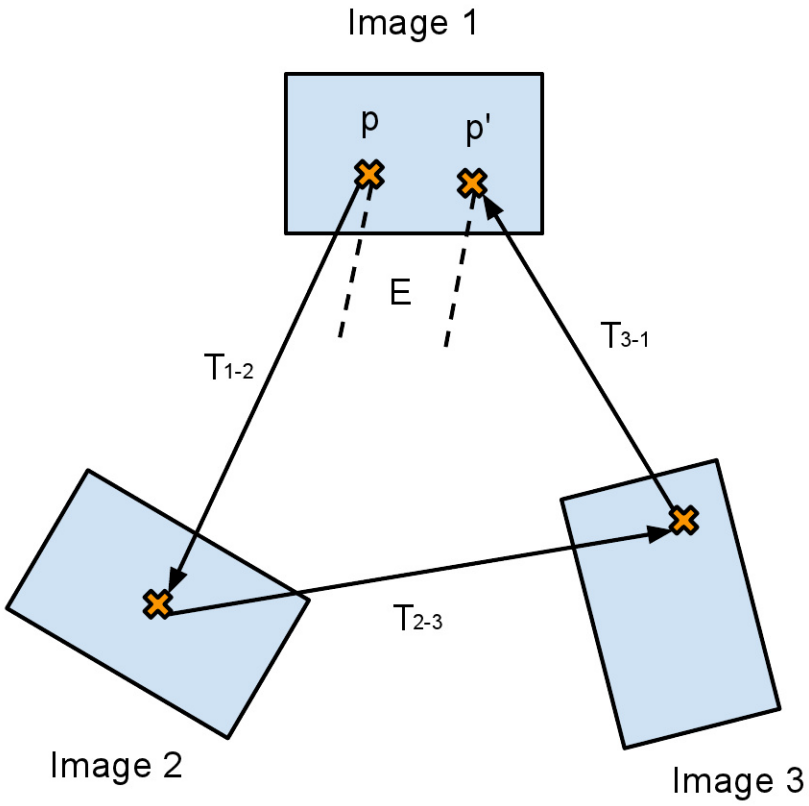


Figure 4.5: Consistency check on three images. The crosses indicate the corresponding locations between all the three images when point p is transformed to image 2, to image 3 and then back to image 1. Therefore, when there are registration errors, the transformations $T_{3-1} \cdot T_{2-3} \cdot T_{1-2}$ will map the point p back into its own coordinate system but in a different position p' (with a displacement error E).

For each dataset of ten images, the registration has been independently applied to three subsets (subset 1 contains images 1, 2, 3, 4, 5; subset 2 contains images 4, 5, 6, 7, 8; subset 3 contains images 7, 8, 9, 10 ,1). Therefore, although the transformations within each subset will be consistent, by using transformations from each independent subset in the loop, it is still able to calculate a consistency value. Due to the overlap of two images between subset 1 and 2, two images between subset 2 and 3, and one image between subset 3 and 1, I am able to calculate four different loops around the three subsets and these loops are $[T_{7-1} \cdot T_{4-7} \cdot T_{1-4}]$, $[T_{7-1} \cdot T_{5-7} \cdot T_{1-5}]$, $[T_{8-1} \cdot T_{4-8} \cdot T_{1-4}]$, $[T_{8-1} \cdot T_{5-8} \cdot T_{1-5}]$. An anatomically relevant region of interest was selected in image 1 with size of $100 \times 100 \times 100$ voxels (in total 1,000,000 voxel points), which is centred on the septum. After mapping those voxels positions using consistency checks, the mean displacement error (MDE) was calculated and averaged from the four different loops for each of the the three volunteer datasets.

Method for capture range experiment

Another important performance value for a registration algorithm is capture range. In order to find out the capture range of registration algorithm, the standard evaluation method proposed by van de Kraats et al. (2005) was used. The first step is to create the starting positions. For each datasets, random errors have been added to the six degrees of freedom of the registration transformations from T_{1-2} to T_{1-10}, to be T'_{1-2} to T'_{1-10}. The same region of interest (as used in consistency check) was used to calculate the initial mean displacement error $IMDE = \frac{1}{M}\sum_{j=1}^{M}(T' \cdot p - T \cdot p)$ for

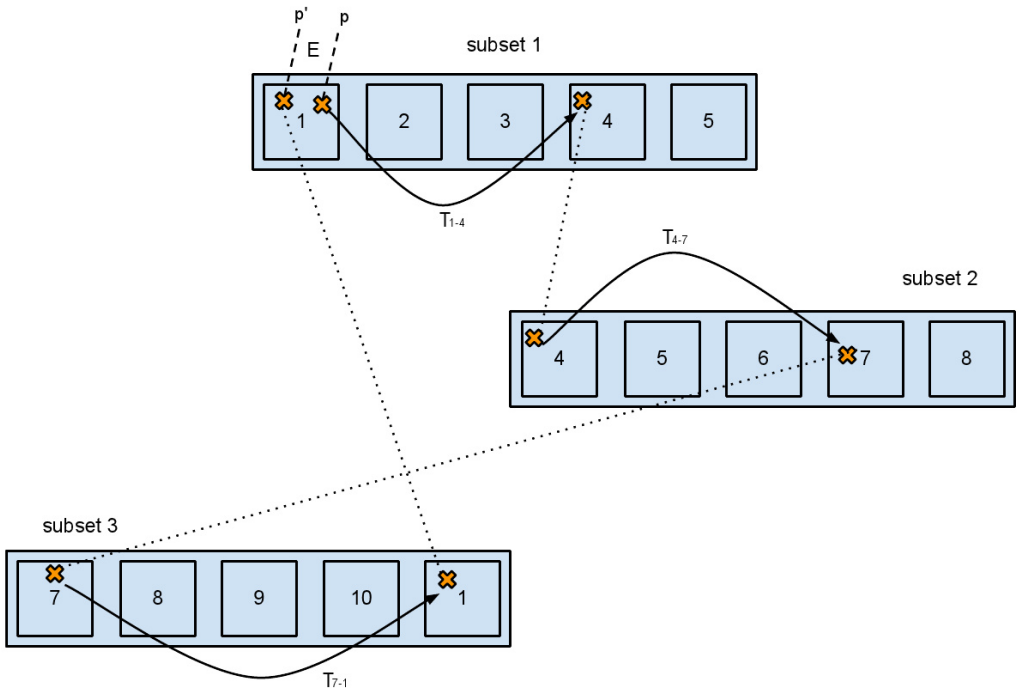


Figure 4.6: Consistency check on semi-simultaneous registration. The registration has been independently carried out on three subsets (1, 2, 3) of the entire image dataset. The crosses indicate the corresponding locations between images when applied with transformations, e.g. T_{1-4}, T_{4-7}, T_{7-1}. If there is registration error, the $T_{7-1} \cdot T_{4-7} \cdot T_{1-4}$ will map the point p in image 1 back to its own coordinate system but in a different position p' (with a displacement error E).

each of the images from I_2 to I_{10}. Fifteen intervals were chosen for the IMDE (0-1, 1-2, ... 14-15 mm) and starting positions were generated until 10 choices for each of the intervals for each images from I_2 to I_{10} were obtained. The second step is to carry out repeated registrations using the pre-defined corresponding starting positions from each interval, and calculate the final mean displacement error $FMDE = \frac{1}{M}\sum_{j=1}^{M}(T'' \cdot p - T \cdot p)$ for each of the images, where T'' is the transformation after registration. In total 450 registrations (3 datasets $\times$ 15 intervals $\times$ 10 starting positions) were carried out. The success criteria for registration is defined as when the value of $FMDE$ for each of the images in a dataset is less than the minimum stepsize of $0.0625\,mm$ for registration algorithm (as described in Section 4.3). The specific method used to define the capture range will be presented and discussed along with the results in Section 4.5.

4.4.3 Compounded image quality validation methods

Most of the previous compounding validations have used SNR and contrast for the numerical image quality indexes. The final compounded images were analysed using both visual inspection and by calculating SNR and contrast values on three different features: left ventricle lateral wall, ventricular septum and anterior mitral valve annulus. These three features are clinically relevant structures which are difficult to see in certain clinical cases from the original images, and can be useful to define the endocardial border of left ventricle. Therefore, improving the quality on these positions should also aid the performance of post-processing algorithms, e.g. segmentation of left ventricle. Six points were manually picked on each feature

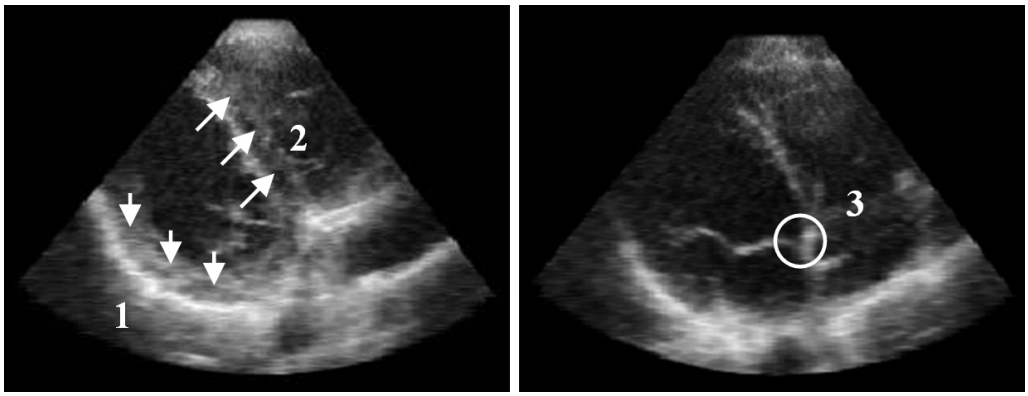


Figure 4.7: Manually defined positions used to define regions where SNR and contrast are calculated. Feature 1 is left ventricle lateral wall; Feature 2 is ventricular septum; Feature 3 is anterior mitral valve annulus. (Dataset-Q1 is shown here for demonstration.

from the original images by experienced echocardiographers (Figure 4.7). A series of background points were also picked within the blood pool of the left ventricle (in original images) so that they were approximately within 10mm of each picked feature point. Regions of $3 \times 3 \times 3$ voxels around each point were used in the following analysis.

We define the SNR in the standard way i.e. $SNR = \frac{Signal}{Noise}$, where $Signal$ is the mean value of the intensities around the manually defined points on the features, and $Noise$ is the standard deviation of intensities around the points picked in the blood pool next to each feature. The contrast is defined as $C = Signal - Background$, where $Background$ is the mean value of intensities around the blood pool points. Final results are averaged SNR and contrast values calculated over all the points picked on each feature.

Table 4.1: The result of consistency experiment on semi-simultaneous registration from three volunteer datasets. MDE stands for mean displacement error.

Dataset	MDE (mm)
Dataset-Q1	1.267
Dataset-Q2	1.085
Dataset-Q3	1.547

4.5 Results

The result of the semi-simultaneous registration consistency check on three volunteer datasets is shown in Table 4.1. The average value of 1.3 mm ($< 2\,mm$) mean displacement error is considered to be very good for echo image registration as reported by the authors of the phase-based similarity measure (Grau et al., 2007).

The percentage of successful registration is shown in Figure 4.8. The graph shows 100% success up until $IMDE$ exceeds $6\,mm$, and then (up to $10\,mm$) initially a slow and then a much faster reduction in success rate occurs as $IMDE$ increases. The required capture range depends on application. If the capture range is defined as the 95% success range as proposed by van de Kraats et al. (2005), then the current registration has a capture range of $9\,mm$. However, if an automatic system requires an 100% successful rate, then the capture range will be defined as $6\,mm$.

The results from the compounding process are presented in Figures 4.9 and 4.10. The results from all three datasets were similar in terms of the trend with respect to increasing the number of images, and the relative performance of the different compounding methods. However the actual numerical values are not comparable due to anatomical variation and dif-

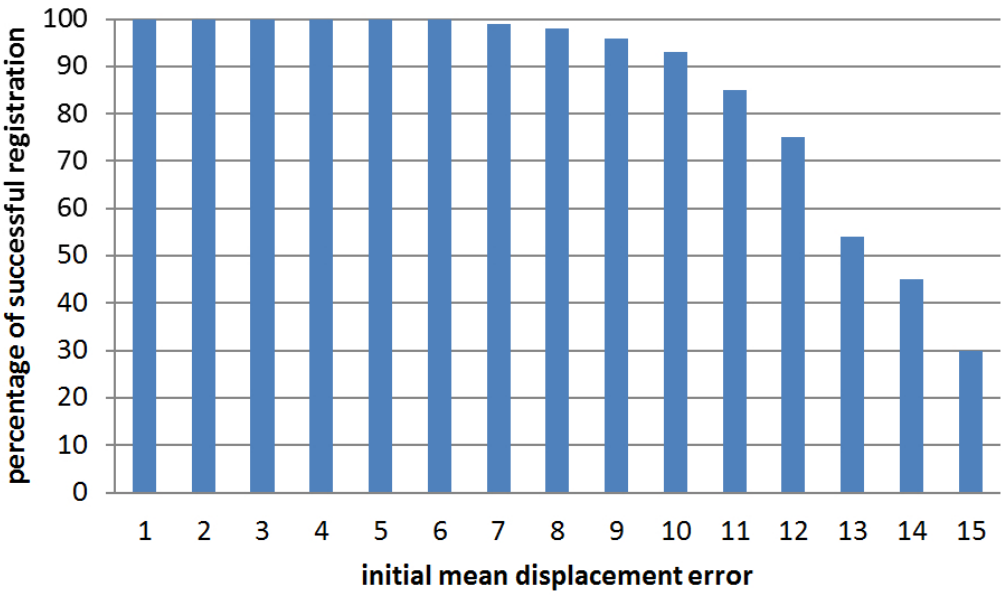


Figure 4.8: The percentage of successful registration on semi-simultaneous strategy.

ferent manual setting adjustments on the scanner, so presenting results averaged over all datasets is not useful. Instead, for clarity, I only present graphs for Dataset-Q2. The numerical results (Figure 4.9) are presented in the form of graphs which show how SNR and contrast vary, for each compounding method, as the number of images increases. As expected, the graphs show the mean method performs well in terms of SNR, but contrast is unchanged. The opposite occurs when using the maximum method: the contrast improves however SNR remain roughly constant. There is a general increase in SNR as the number of images increases for all methods except the maximum. The Phase selection methods performed well both in terms of SNR and contrast.

Reformatted 2D slices through an original single 3D volume and the compounded 3D echo volume from Dataset-Q3 are presented in Figure 4.10. The single volume was randomly chosen for demonstration, and is compared with compounded images produced by applying different compounding methods

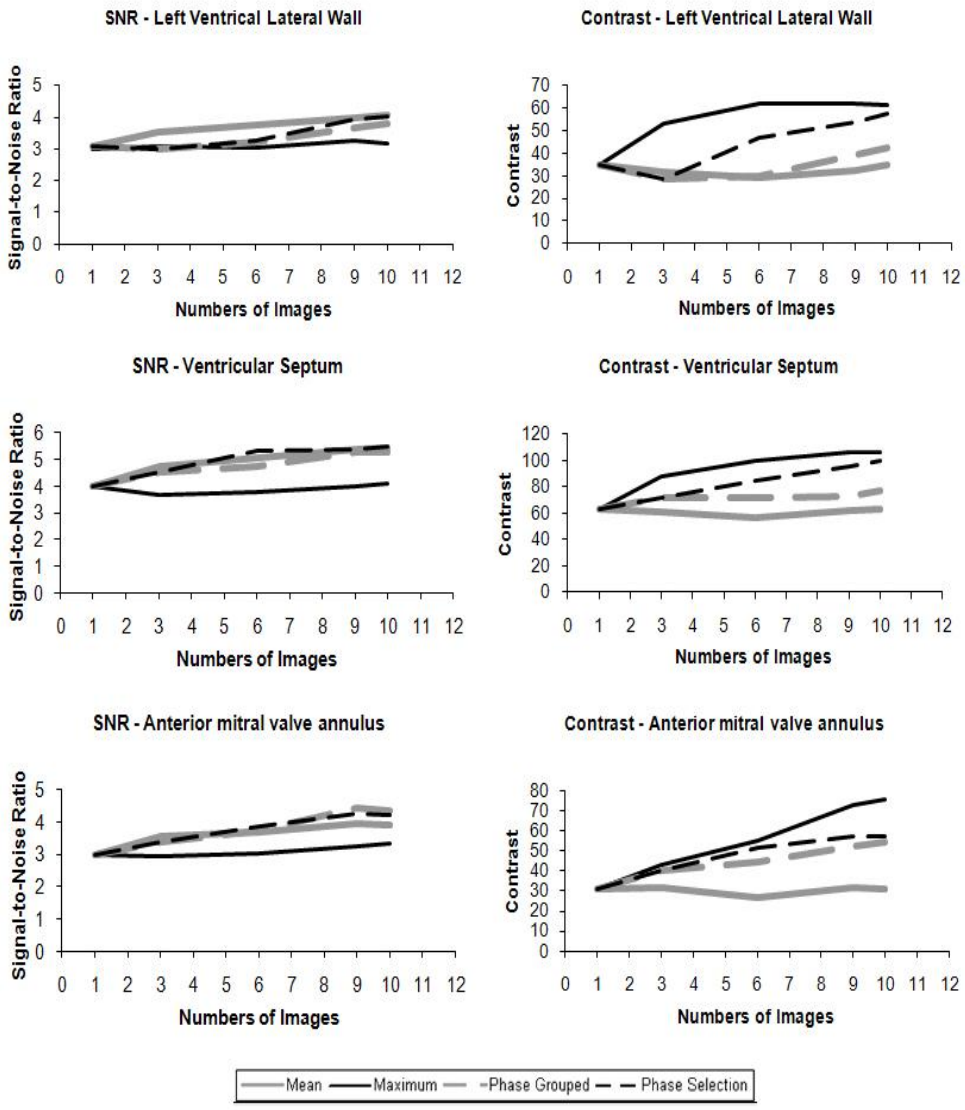


Figure 4.9: The SNR (left column) and contrast (right column) on three features (left ventricle lateral wall, ventricular septum and anterior mitral valve annulus) when compounding with increased number of images using different compounding methods (Dataset-Q2 is shown here for demonstration).

(using 5 images and 10 images compounding). The mean compounded image can be observed having less noise but looks blurred. The maximum compounded image has very good contrast, but has higher intensity noise. The phase grouped and phase selection techniques both maintain the contrast of image features and have lower noise than the maximum image. The phase selection image has better contrast than the phase grouped image.

Furthermore, compounding using 10 images performs better than using smaller number of images such as 5. Evidence from numerical results (Figure 4.9) show positive slope are presented when compounding 10 images compared to 5 (excluding the maximum method for SNR and the mean method for contrast). Visual results (Figure 4.10) also show using more images improves image quality.

4.6 Discussion

The results in Figure 4.9 show evidence that increasing the number of images can improve SNR and contrast but this depends on the compounding strategies used. This was an expected result and has previously been shown for 2D B-mode ultrasound by Rohling et al. (1997). The four compounding methods have performed differently as shown numerically in Figure 4.9 and visually in Figure 4.10. The Phase Grouped method uses all the image data for each voxel and relies on the compounding process to filter out the less useful intensities. Therefore, it should have a higher SNR than the Phase Selection method. But repeatedly grouping images and combining intensities using a weighted mean results in smoothing the image. The

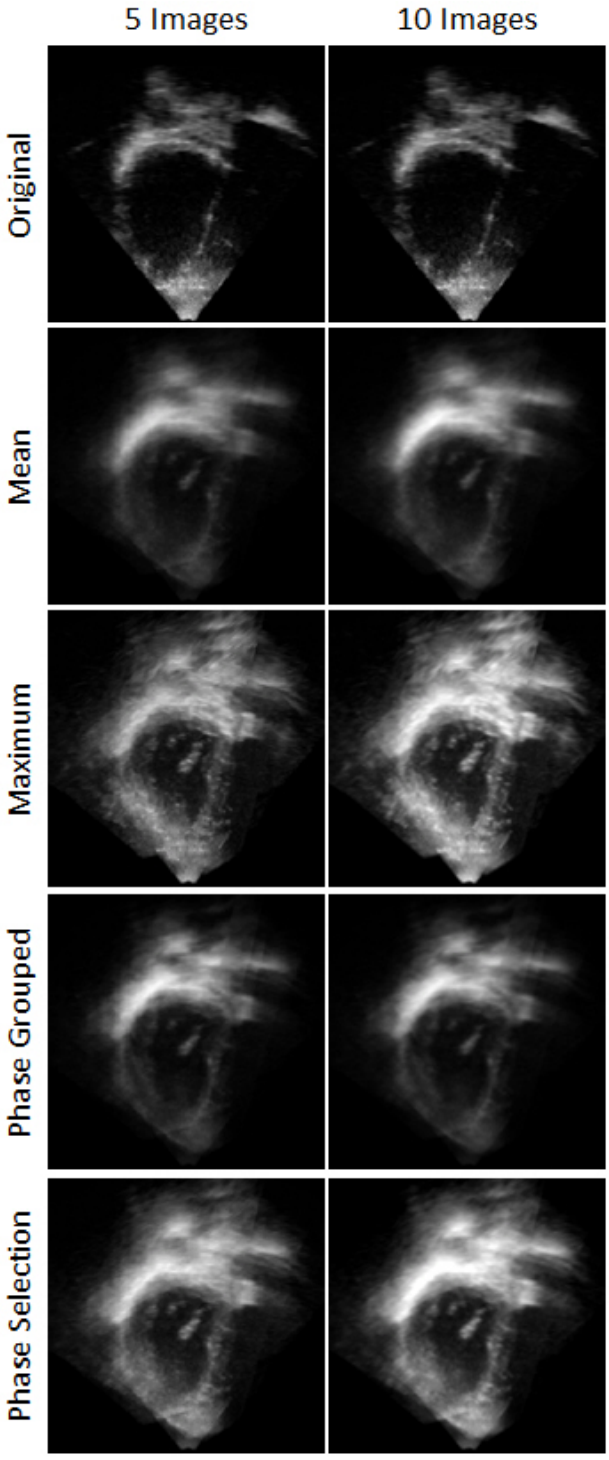


Figure 4.10: Example of 2D slices through an original 3D volume and the final compounded 3D echo images produced by applying different compounding methods. 5 images and 10 images compounding are compared (Dataset-Q3 is shown here for demonstration).

Phase Selection method relies upon an image selection step which is carried out for each voxel position to determine which images are the most suitable for compounding that voxel. This has advantages in terms of less computational processing, and the selection step can be specifically tailored towards particular image features or to reduce the possibility of including echo image artefacts in the final compounded volume. The Phase Selection method I implemented in this work used a very simple selection criterion to choose the three highest intensities at each voxel and pass them through the standard phase-based method. It produces a high contrast image. As the Mean method and the Maximum method performed well in terms of SNR and contrast respectively, it could be good to use both methods as the benchmark in my further investigations later on.

Compounding larger numbers of 3D echo images improved image quality in terms of better SNR and contrast. However, using larger numbers of images means that the total time taken for data acquisition, image registration and compounding processes will be longer. Real-time imaging could be an option to rapidly acquire larger numbers of images. Currently, my compounding method uses rigid registration to align the images. Small registration errors are accumulated when combining larger numbers of datasets. This will result in some structures appearing to be thicker than their real size, blurring of anatomical structure, or misregistration of fast moving structures such as the mitral valve. The final result depends on whether the compounding methods can filter out these registration errors.

The automatic registration method used was semi-simultaneous strategy with a phase-based similarity measure, and a starting position found using

optical tracking. I observed problems with this method if the initial alignment is far from the optimal solution. In these cases the algorithm either did not iterate to the optimum solution, or it took an excessive amount of time. This was caused, I believe, because as a given image is being registered it can receive contradictory information from the other images in the data set as to where the correct registration position lies. Therefore, in the next section a modified semi-simultaneous strategy will be proposed which aims to improve the registration robustness.

Due to anatomical variation and the echocardiographer being allowed to alter the gain setting on the echo machine, image intensities for a given piece of tissue varied between different acquisitions. This could make it difficult for a compounding algorithm to filter the best intensity information. For example, if an intensity value could be passed through to the final image because the clinician set a high gain value for that image rather than it being the most representative intensity. Therefore, investigations will be carried out in the next chapter to produce a post-processing protocol which can normalise the intensity of acquisitions in order to avoid significant intensity variation.

In image compounding I am attempting to improve the quality of the echo images, in terms of reducing noise, improving contrast and increasing the field-of-view. However, quantitatively assessing image quality is a very difficult task, and SNR and contrast values only capture particular aspects of image quality. It is also very difficult to define what the target compounded echo image should actually look like. Compounding is not about simply averaging image noise, but also includes boundary enhancement and arte-

fact removal. In addition the final objective of what the compounded image should look like will also dependent on the final use of the image. More detailed discussion and definition of our final compounded image model will be described in the next chapter.

4.7 Conclusion

In this chapter, I have investigated compounding using increased numbers of 3D echo images. Prior compounding methods like Mean, Maximum, and extended Phased-based methods were compared to see how image quality is improved by using more images. Results show using more images are useful for averaging noise to increase the SNR for the Mean and Phase Grouped compounding methods, and for contrast improvement for the Maximum and Phase Selection compounding methods. As some parts of the heart can only be imaged well in certain view angles, in order to construct a more coherent final image, information from a larger numbers of multi-view images will help to reduce the angular dependent effects.

If a decision is made that a larger numbers of images will be used, then this can provide the opportunity to produce novel compounding techniques. For example, larger numbers of images could allow detection of artefacts (as outliers) and reducing the effect of registration errors, which will be further investigated in the next chapter.

Chapter 5

Echo compounding using feature consistency

This chapter presents a novel compounding method for larger numbers of multi-view 3D echocardiography volumes. I initially explain the motivation behind my compounding technique. I then define an ideal image which my compounding algorithm is aiming to produce. My algorithm is specifically designed to address areas containing artefacts, and to work with larger numbers of images. In order to further improve compounding, I will also investigate strategies to align larger numbers of images and methods which can cope with global brightness variations between images.

5.1 Motivation

As discussed in Chapter 3, previous 3D compounding methods have typically been applied to a relatively small number of images, and none of them have been specifically designed to work with larger numbers of 3D echo im-

ages. In Chapter 4 I investigated the effect of using an increased number of 3D echo images for compounding. The results indicated that compounding more images can further improve image SNR and/or contrast depending on the specific compounding strategy used.

However, there are specific problems associated with compounding larger numbers of images. For example, using more images could bring in more registration errors to the compounding process, causing blurring for example. One method to reduce these possible effects would be have a compounded algorithm which has a mechanism of filtering registration errors. Also, global brightness variations have been observed between acquired multi-view images due to a number of reasons such as manual adjusting the overall gain setting on the echo machine. Both registration errors and intensity variations will reduce the ability of compounding algorithm to improve the image quality.

Artefacts are common in echo images. The presence of artefacts either causes signal loss along the beam direction (due to ribs or air in the lung blocking the sound wave), or introduces false information in the images (e.g. reverberation artefacts described in Section 2.1.4). These artefacts can be a major factor affecting echo image quality. Surprisingly none of the previous compounding methods have directly addressed the issue of echo artefacts. My compounding method will directly address issues caused by artefacts.

Furthermore, none of the previous compounding work have explicitly defined what their ideal final compounded image would look like i.e. under ideal circumstances what their proposed compounding process is attempting to achieve. In this chapter, my definition of an ideal compounded echo image

will be proposed, and I explain how this definition has influenced the design of my proposed compounding algorithm.

5.2 Proposed ideal compounded image

The reason for defining an ideal compounded image is to have a target to aid algorithm design. Although there could be many definitions of an ideal image, I have chosen to base my definition on the underlying ultrasound physics (i.e. how ultrasound interacts at positions where changes in acoustic impedance occur). I propose that an ideal echo image should present high quality anatomical structures with high SNR and contrast, and no artefact. In order to define my ideal compounded image I consider three types of regions within the image. The regions chosen do not relate to specific anatomy, but can be used to represent nearly all the anatomy within the human body relevant to ultrasound examination. For example specific anatomy such as papillary muscles and the endocardium are all included in my definition of an ideal image as soft tissue structures, which will have reflection effect on boundaries between different tissue type and will have a speckle pattern within homogenous tissue regions.

boundary region : where there exists a boundary between two tissue types which have different acoustic impedance values. In these regions the intensity of the reflected ultrasound signal will primarily be a function of two variables: the change in acoustic impedance ΔZ and the angle between the beam and boundary surface normal α. I propose that the intensity in my "ideal" compounded image for a boundary

region should be as if the region were insonated with a ultrasound beam which is normal to the boundary i.e. $\alpha = 0°$. This would remove any angular dependence on compounded image intensity, and so the intensity in these regions would be purely a function of ΔZ. The aim is to remove acquisition dependent effects (insonation angle) and result in a final image which is just dependent on the underlying anatomy. The Incidence Angle Weighting method (Leotta and Martin, 1999) and the Phase-based method (Grau and Noble, 2005) used this idea to assign high weighting values to those regions with less angular difference between the beam and boundary surface normal.

homogeneous tissue region : in this region the reflected ultrasound signal is caused by constructive and destructive interference between waves reflected from the numerous small scale structures within homogeneous tissue. This creates the characteristic speckle pattern observed in echo images. In homogeneous tissue I model the intensity of the reflected ultrasound as a function of a single variable: tissue scattering property. I propose that the "ideal" image should have uniform intensity, which is the mean of the speckle intensity distribution of that region. All the previous methods (except the Maximum method) have used a form of averaging in homogeneous regions, which averages out speckle unless coherent speckle is present in all images.

an artefact region : can occur due to a variety of reasons. Artefacts are commonly caused by strongly reflecting boundaries or poor probe-patient contact. The resulting artefacts have low intensity values for

all voxels along the beam path behind the strongly reflecting boundary. These are included in my imaging model using a function representing a set of positions along the beam path from the transducer to the specific region. Artefacts are not directly related to a specific region of the body as the previous two regions were, but are simply regions within the echo image that contain artefacts. I propose that artefact regions should have no contribution to the final intensities in my "ideal" compounded image.

5.3 Consistency based compounding method

This section details my proposed compounding method. It employs a technique designed to work with larger numbers of images and in particular addresses echo artefacts. As shown in Figure 5.1, if there is a strong reflecting structure (e.g. rib) which causes a shadowing artefact, then the position of the artefact (A1 and A2) will vary greatly (with respect to the underlying anatomy) if different transducer positions (P1 and P2) are used. Therefore, if a large number of images from different views are available, it may be possible to detect artefact regions as they will appear at different places while the underlying anatomy stays consistent in a set of registered multi-view images. In this case, artefects could be thought of as outlier region within the set of images. Therefore, in this section, a compounding technique based on feature consistency is proposed, which takes the advantage of multi-view images and directly tackles the issue of echo image artefacts.

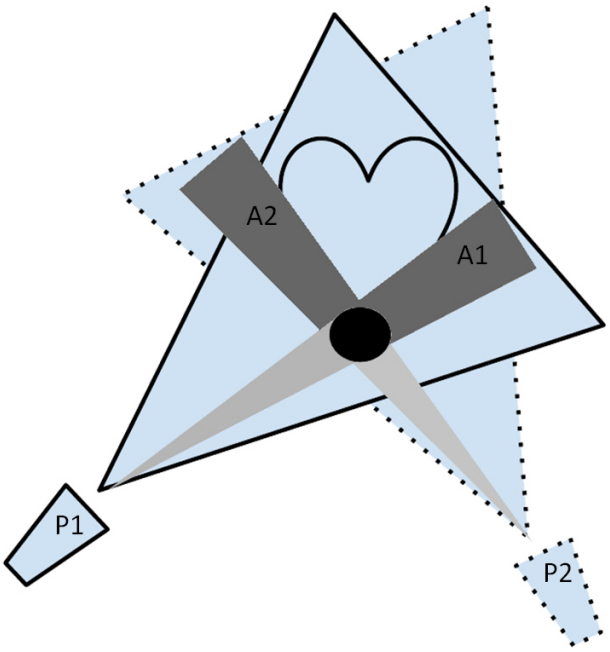


Figure 5.1: Demonstrating the changes on artefact regions between different views. The black dot present a strongly reflecting structure (e.g. rib). By taking images from two different view angles (P1 and P2) the anatomical feature should stay consistent between views while the artefacts (A1 and A2) caused by the strongly reflecting structure vary greatly in position with respect to view angle.

5.3.1 Compounding based on feature consistency

I propose a compounding method aimed at reducing the noise in homogeneous regions, enhancing contrast at boundaries and removing the effect of artefacts. The algorithm will use a set of N images, and compare small regions of each overlapping image at each voxel location. The final compounded image is represented by I_f and is produced by a weighted mean scheme

$$I_f(\mathbf{x}) = \frac{\sum_{i=1}^{N} w_i(\mathbf{x}) * I_i'(\mathbf{x})}{\sum_{i=1}^{N} w_i(\mathbf{x})} \tag{5.1}$$

where $w_i(\mathbf{x})$ is the weighting factor for each voxel location, $\mathbf{x}$, in each image, i. The detailed information of how to obtain the weights, $w_i(\mathbf{x})$, is explained in the following paragraphs.

In both boundary and homogeneous tissue regions the following observation is used to reduce the effect of artefacts: intensities within boundary and homogeneous regions depend primarily on the tissue within a local region, whereas the artefact is dependent on all the voxels along the beam path. Therefore, the position of artefact regions should vary to a much greater degree between different views than non-artefact regions, as the set of voxels in the beam path will vary greatly with respect to view angle.

The main novelty of my technique is to make use of this observation in the following way: a region in image i, $R_{\mathbf{x}i}$ will be compared to the registered regions in the other images, $R_{\mathbf{x}j}, j = 1, \ldots, N, j \neq i$ using a consistency measure. As artefact regions are more dependent on view direction, there should be lower consistency value if region $R_{\mathbf{x}i}$ contains artefact, compared

to if region $R_{\mathbf{x}i}$ contains a boundary or homogeneous tissue. The weighting, w_i in equation (5.1) is calculated from my consistency measure, and I use different measures for boundary and homogeneous tissue regions, which are now described in detail.

As a measure of consistency at boundaries I use normalized cross correlation (NCC):

$$NCC_{ij}(\mathbf{x}) = \frac{\sum_{\mathbf{y}\in R_{\mathbf{x}}}(I'_i(\mathbf{y}) - \bar{I}'_i)(I'_j(\mathbf{y}) - \bar{I}'_j)}{(n-1)\sigma_i\sigma_j} \tag{5.2}$$

where $\bar{I}'$ represents mean intensity and σ represents standard deviation of intensity in region $R_{\mathbf{x}}$ which contains n voxels.

For a given region $R_{\mathbf{x}}$ each image is taken in turn, and the NCC between this image and all the other images is calculated (Figure 5.3). A boundary weighting term, $w1(\mathbf{x})$, is calculated by averaging NCC values over all the images. If region $R_{\mathbf{x}}$ contains a boundary, I would expect a high value of $w1$ (as shown in Figure 5.4 (b,d)). Note that the intensity at a boundary depends on the angle between the boundary surface normal and ultrasound beam. However, if region $R_{\mathbf{x}}$ is small enough the boundary can be considered to be a flat plane and so there should still be a linear relationship between image intensities from different view directions (Figure 5.2). Therefore, NCC was chosen as the boundary consistency measure as it should be invariant to this angle effect and will produce a high values at boundaries even between different view directions. If region $R_{\mathbf{x}}$ contains an artefact in image i, I would expect there to be low NCC values, resulting in a low value for $w1(\mathbf{x})$.

Although $w1(\mathbf{x})$ provides a high value at a boundary region, it contains

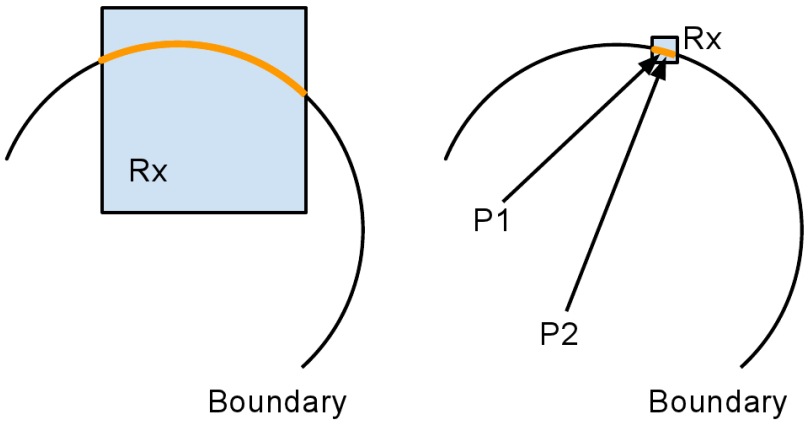


Figure 5.2: Demonstrating a region $R_{\mathbf{x}}$ on boundary. Right diagram shows when the $R_{\mathbf{x}}$ is small enough, the curved boundary can be considered as a flat plane, compared with the left diagram. By sending ultrasound beam from two different view angles (P1 and P2), the amplitude of reflected ultrasound from the small flat plane will only depend on the angle of incident. Therefore, there should be a linear relationship between image intensities from different view directions for the small enough region.

no information on the magnitude of the boundary reflection. To approach the ideal image we defined in Section 5.2, i.e. to reduce the angular dependency of boundary reflections and to obtain an intensity value as if the boundary was normal to the beam direction, the maximum intensity value should be used from the image set at position $\mathbf{x}$. However, this will result in high intensity noise in the final compounded image. Therefore we introduce another factor, which we call boundary boosting $s_i(\mathbf{x})$,

$$s_i(\mathbf{x}) = \left(\frac{I'_i(\mathbf{x})}{I'_{max}(\mathbf{x})} \right)^q \tag{5.3}$$

where q is used to adjust the strength of the boosting, and $I'_{max}(\mathbf{x})$ is the maximum intensity value in a dataset. The effect of the boundary boosting term is to preserve high intensity edge reflections in the compounded images, while reducing the impact of lower intensity reflections due to lower incident angles. It would be also useful in the situation where a large percentage

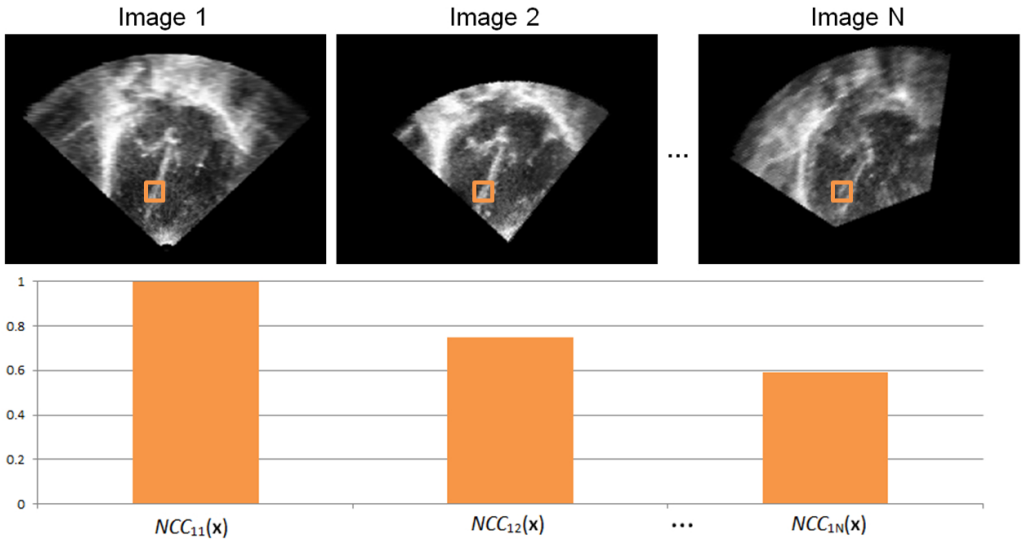


Figure 5.3: Illustration of how boundary consistency weights, $w1(\mathbf{x})$, are calculated. Top row depicts a set of registered echo images from different view directions. Weights, $w1(\mathbf{x})$, are calculated for each voxel, by comparing image information in a local region around that voxel (one such local region is illustrated by the square in the images). The boundary consistency measure used between regions is normalized cross correlation (NCC). The above graph shows NCC calculated for the local region between image 1 and all the other images in the dataset. The mean of these values is my weight, $w1(\mathbf{x})$, for image 1 at position $\mathbf{x}$. The above is repeated for each image to calculate a complete set of weights at position $\mathbf{x}$.

of signal energy has been lost, however some signal is still detected from a strong edge feature. $s(\mathbf{x})$ is then multiplied by $w1(\mathbf{x})$,

$$w1s_i(\mathbf{x}) = \alpha * w1_i(\mathbf{x}) * s_i(\mathbf{x}) \tag{5.4}$$

where α is a normalisation term, which is a function of $\mathbf{x}$:

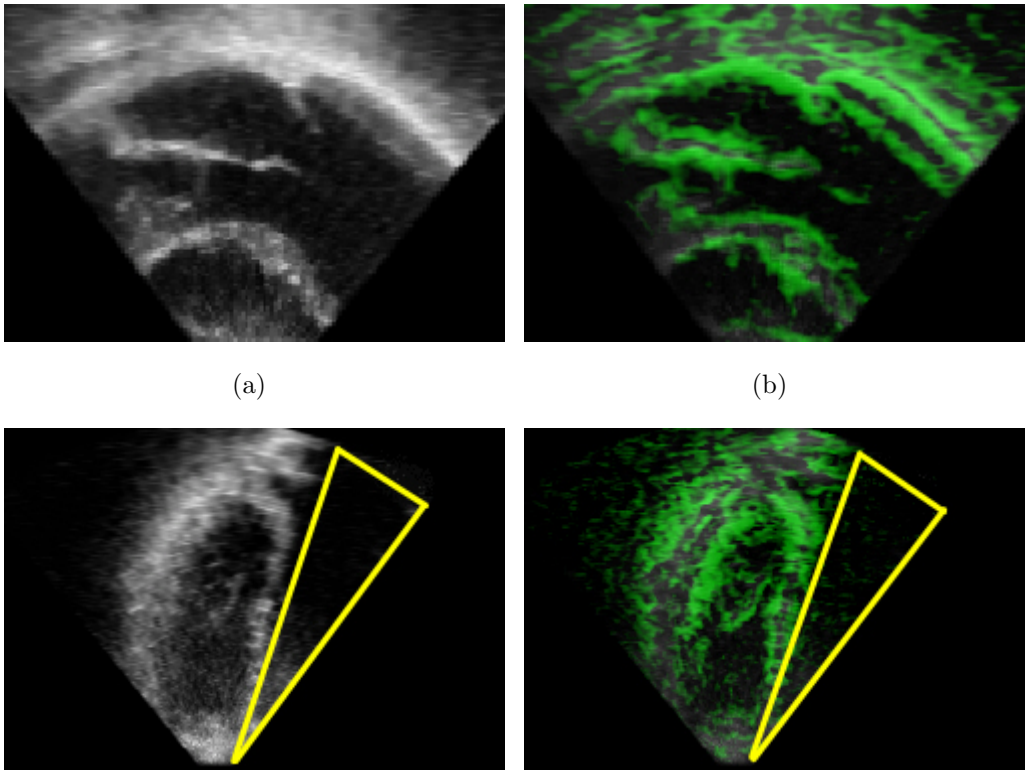
$$\alpha(\mathbf{x}) = \frac{\sum_{i=1}^{N} w1_i(\mathbf{x})}{\sum_{i=1}^{N} w1_i(\mathbf{x}) * s_i(\mathbf{x})} \tag{5.5}$$

In a homogeneous region, as it contains speckle which is not coherent between different view angles, I also expect a low $w1(\mathbf{x})$ value. To distinguish between homogeneous and artefact regions I calculate a value H which returns a high value if two regions have similar standard deviations of intensities:

$$H_{ij}(\mathbf{x}) = \frac{u}{|\sigma_i(\mathbf{x}) - \sigma_j(\mathbf{x})| + u} \tag{5.6}$$

where $\sigma(\mathbf{x})$ is the standard deviation of intensity values in region $R_{\mathbf{x}}$, and u is used as a sensitivity measure ($u = 1$ was used in this thesis). A homogeneous region weighting term, $w2(\mathbf{x})$, is calculated by averaging H values over all the images. Previous work has also used a high order statistical method (skewness) to differentiate artefact and speckle (Drukker and Giger, 2003). I expect a high H value if region $R_{\mathbf{x}}$ contains homogeneous tissue (as shown in Figure 5.5 (b,d)).

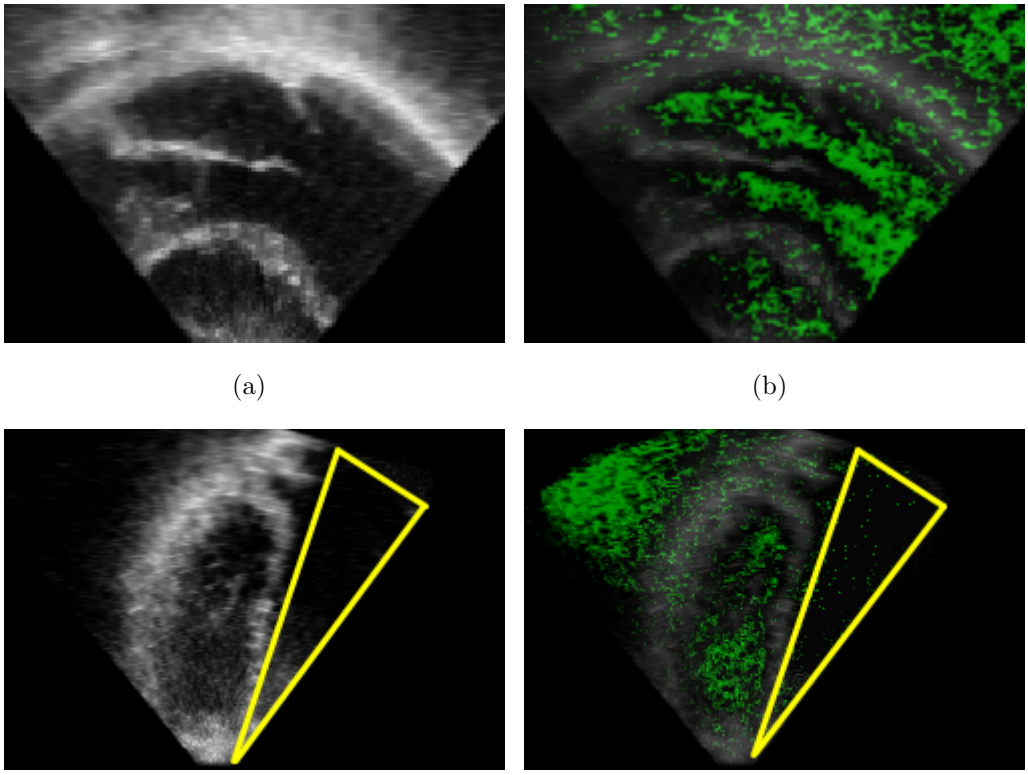
Therefore, $w1s(\mathbf{x})$ is used to locate boundaries and enhance them to improve contrast, and $w2(\mathbf{x})$ is used in homogeneous regions to increase SNR. The $w_i(\mathbf{x}) = w1s_i(\mathbf{x}) + w2_i(\mathbf{x})$ is the overall weight, w_i, which is used to produce the final compounded image $I_f(\mathbf{x})$ as defined in equation (5.1) at the beginning of this section.



(a) (b)

(c) (d)

Figure 5.4: Example of regions highlighted by weight $w1$(weighting values are normalized between 0 to 1). (a) single original image slice showing the parasternal long axis view; (b) $w1$ was calculated and overlayed with the original image showing boundaries (the "double edges" on the valve occurs when there are two step edges i.e. the changes of acoustic impedance happens; (c,d) are presented in the same way as (a,b), but with an artefact region (due to shadowing) labeled with triangle. As expected, $w1$ shows low weighting values inside the artefact region.



(a) (b)

(c) (d)

Figure 5.5: Example of regions highlighted by weight $w2$ (weighting values are normalized between 0 to 1). (a) single original image slice showing the parasternal long axis view; (b) $w2$ was calculated and overlayed with the original image showing the homogeneous regions; (c,d) are presented in the same way as (a,b), but with an artefact region (due to shadowing) labeled with triangle. As expected, $w2$ also shows low weighting values inside the artefact region.

5.3.2 Empirical parameter estimation

The values of $R_{\mathbf{x}}$, q and u were all determined empirically in order to find the right parameters towards the best results of SNR and contrast improvement on image quality. Only a small subset (3 datasets) from the volunteer datasets was used to avoid bias. Values for $R_{\mathbf{x}} = 3 \times 3 \times 3$, $5 \times 5 \times 5$ and $7 \times 7 \times 7$, $q = 1, 2, 3$ and $u = 1, 2, 3$ were investigated. It was found that $R_{\mathbf{x}} = 5 \times 5 \times 5$, $q = 2$ and $u = 1$ provided the best results. These parameters were then held constant for all the subsequent experiments.

5.4 Additional requirements for 3D echo compounding

A complete system to produce compounded images can be separated out into a number of steps. I have split the process up into three steps. In order to carry out compounding, the system requires oversampled data. The first step is "data acquisition" which includes any additional requirements or additional data capture necessary for the subsequent image processing. To use these multiple datasets in the compounding process the system needs to be able to map spatial positions and cardiac phases from one image set to all the others (which overlap). This process is termed "image registration" and is the second step required by my system. The third step is the actual "image compounding", the process by which an algorithm has as input a set of intensity values from a number of overlapping datasets and these are processed to a single final value which will represent the tissue at that spatial/temporal location in the output compounded image. However, issues

aligning larger numbers of images and global intensity variations between images have been raised in Chapter 4. Therefore, in this section, I will concentrate to further improve the registration for larger numbers of images and to cope with global brightness variations between images. They are important for a better 3D echo compounding.

5.4.1 RCF group-wise registration

As discussed in Chapter 4, when the initial alignment is far from the optimal solution, the registration algorithm may not iterate to the optimum solution. Therefore, I have slightly modified the previous implementation of semi-simultaneous (Section 4.3) so the method registers the closest views first (which is denoted by RCF and as shown in Figure 5.6). The RCF method uses the starting position information (obtained from optical tracking in most of our experiments) to calculate which two images are closest with respect to view angles. These images are registered together using the previous implementation of semi-simultaneous strategy. After aligning the first two images, then the image which has the minimum average view angle to the two registered images is included into the registration process. The algorithm then registers all three images together, picking one image at a time and aligning it to the two other images averaging the value of the similarity measure to both images. When no improvement in the similarity value can be achieved the algorithm will include a fourth image, chosen by the closest average view angle to the registered images, and this process is repeated until all the images are registered.

To assess the effect of these changes on registration consistency and

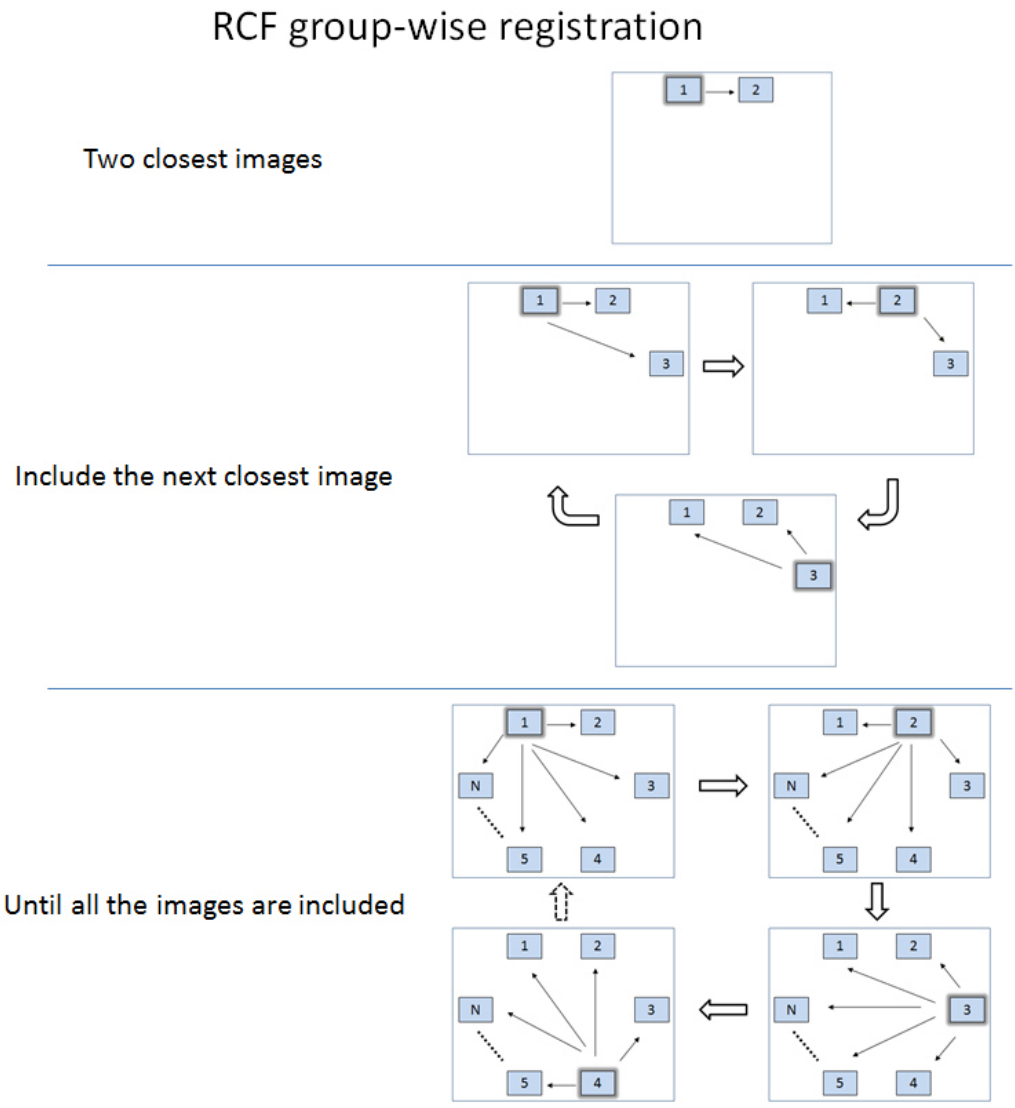


Figure 5.6: Illustration of *RCF* group-wise registration. The method uses the previous implementation of semi-simultaneous strategy in each of the following steps: It starts with aligning two images with closest view angle (Image 1 and 2 in this case). Then the third image (Image 3) which has the minimum average view angle to the two registered images is included into the process until no improvement in the similarity value can be achieved. This process is repeated until all the images are included and registered.

Table 5.1: The result of consistency checks on RCF registration from three volunteer datasets. MDE stands for mean displacement error. The result from semi-simultaneous registration have been included in brackets for easy comparison.

Dataset	MDE (mm)
Dataset-Q1	1.190 (1.267)
Dataset-Q2	1.049 (1.085)
Dataset-Q3	1.445 (1.547)

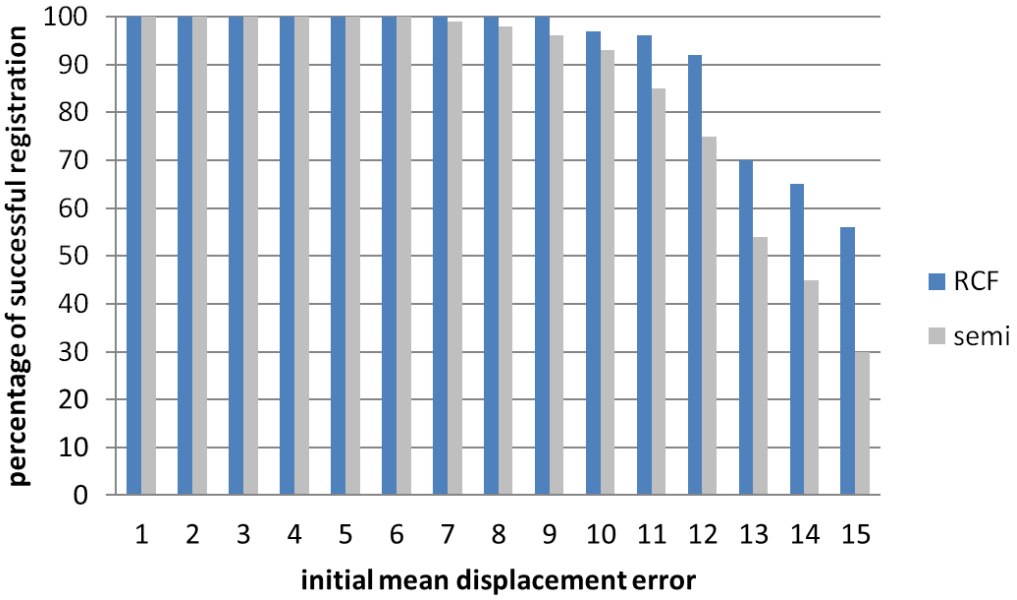


Figure 5.7: The percentage of successful registration on RCF strategy. The result from semi-simultaneous registration (semi) have been included for easy comparison.

capture range, the experiments to determine RCF registration consistency and capture range have been carried out exactly the same as described in Section 4.4.2 using the same datasets and region of interests. Therefore, the results shown in Table 5.1 and Figure 5.7 can be directly compared to the results using the previous semi-simultaneous registration method (Table 4.1 and Figure 4.8) which have been repeated here for easy comparison.

As can be seen from Table 5.1, all the mean displacement errors are smaller when using the new method RCF, however the difference is small

(0.07 mm). As can be seen from Figure 5.7, the RCF method has a much wider capture range 11 mm (compared to 9 mm previously) if 95% successful rate is defined, or 9 mm (compared to 6 mm previously) if 100% successful rate is required. Therefore, for highly automatic system which would require a very high success rate (close to 100%) use of the RCF strategy achieved a 50% improvement on the registration capture range.

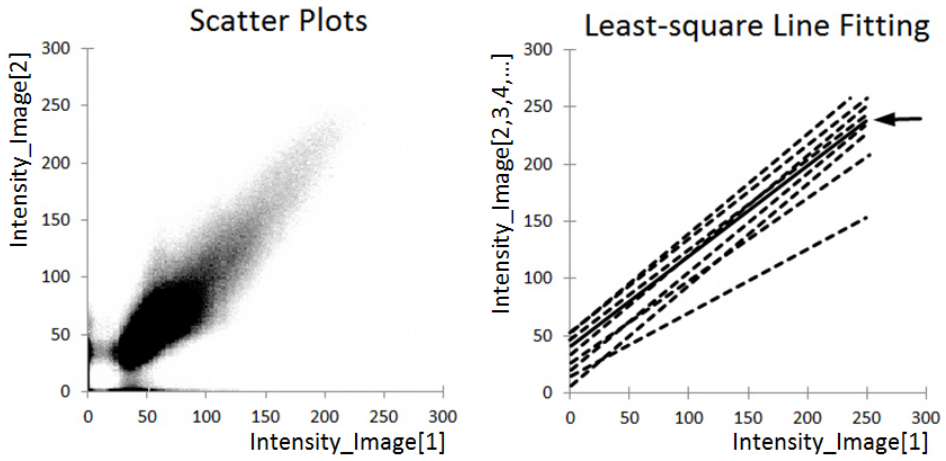
My registration algorithm requires a starting position, which was achieved for most of my experiments by optically tracking the transducer. This allows very accurate tracking of the transducer. The optical tracking system (as described in Section 4.3) is expected to have good accuracy and a latest study (Schmidt et al., 2009) stated a range of value from 0.6 to 29.2 μm, which is well below the manufacturer's standard (up to 0.1 mm). However, errors may come from other factors, one of the largest causes of error in starting position is expected to be variations in heart position due to respiratory motion. The respiratory motion of the heart has a mean value of 4.9 mm with standard deviation 1.9 mm (Shechter et al., 2004). As the position of exhale has been reported to be more reproducible than inhale (Seppenwoolde et al., 2002; Sonke et al., 2008), acquiring data at exhale position could minimize the error coming from the respiratory motion. However, even with respiratory motion error of 5 mm, as long as the effect of other errors (tracking and calibration) can be kept below 4 mm, the RCF algorithm should be able to achieve a high registration success rate.

5.4.2 Intensity normalisation

Differences in the intensity of a particular tissue region between different images can occur due to a number of reasons. These include manual operator brightness changes and the tissue types traversed by the echo signal before reaching the region of interest. These differences are likely to reduce the ability of the compounding algorithm to enhance image quality. No intensity normalization scheme has been reported in previous compounding methods. The following method was proposed to carry out automatic global intensity normalization between echo images in each dataset. Scatter plots are produced relating corresponding intensities between the first echo image in a dataset and each of the other images after registration (Figure 5.8a). A best-fit line (least-square) is produced from each of these scatter plots (Figure 5.8b). These straight lines are defined by a slope and an offset. The image with the median slope is used as a reference image, and the intensities of all the other images are linearly mapped (using the best-fit line) to normalize their intensities with respect to the reference image.

5.5 Conclusion

In this chapter, a definition of an ideal compounded image was proposed and my compounding method was designed to create this ideal compounded image from the individual volumes. I have described a compounding method to take advantages of using larger numbers of images. As artefacts vary between view angles to a greater extent than anatomical features, my consistency based compounding is used to detect artefacts and using



(a) (b)

Figure 5.8: (a) An example of scatter plots between two registered echo images, which shows an approximate linear relationship between image intensities; (b) Result of least-square line fitting between the intensities of first echo image and each of the remaining echo images after registration. The line pointed by the arrow is with the median slope.

more images should provide statistical robustness. Also, in order to further improve the compounding quality, new methods were proposed for image registration and image intensity normalization. Thorough validation on my consistency based compounding algorithm will be carried out in the following chapters using phantom, volunteer and patient data.

Chapter 6

Validation of feature consistency compounding using phantom and volunteer data

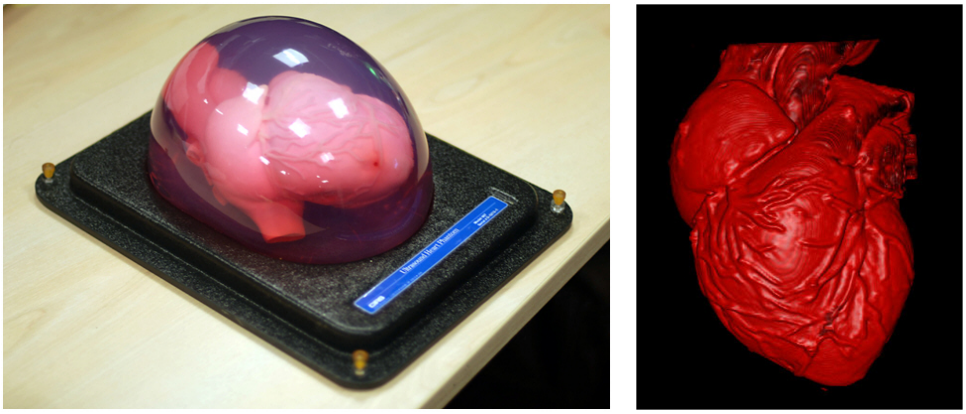
In the literature, previous methods to validate compounding algorithms have used a variety of datasets such as computer simulated images (Grau and Noble, 2005), images of a phantom (Leotta and Martin, 1999; Szmigielski et al., 2010) and/or volunteer/patient images (Leotta and Martin, 1999; Soler et al., 2005a; Grau and Noble, 2005; Rajpoot et al., 2009; Szmigielski et al., 2010). It is relatively easy to obtain a "gold-standard" from simulated data. However, simulated images are open to criticism due to the realism of the simulation.

In this chapter, to assess how my feature consistency based compounding performs, a thorough validation will be carried out using phantom and volunteer data. The experiments are presented separately in the following sections. The value of using those datasets and specific data acquisition pro-

cedures will be described accordingly. My feature consistency compounding method is then compared to the original, uncompounded echocardiography images, and to two basic statistical compounding methods (Mean and Maximum). Image quality improvements using 10 as compared to using 5 images are shown. Both numerical and visual inspection results will be presented and discussed. These validations on phantom and volunteer data will build the foundations for further clinical evaluation in the next chapter.

6.1 Validation using phantom data

In this section, the phantom experiments are described. The advantages of using a phantom are as follows: firstly and foremostly the ultrasound machine is used for image acquisition. This means that the images used should have realistic noise with respect to all the image formation process within the transducer and the echo machine. The realism of the images used will depend heavily on the phantom used. As will be described in more detail below, I have used a commercially produced heart phantom and have increased the clinical realism of the images by placing additional soft and hard tissue structures between the transducer and the phantom to generate a soft tissue layer and artefacts. There are fewer time constraints when imaging a phantom, so that many images can be taken, from lots of view angles (there is no acoustic windows limitations from phantom) and so a "gold-standard" can be obtained. These views, their number and where they were acquired from, can be reduced to increase clinical realism, to be views from standard clinical acoustic windows. However, there are a number



(a) (b)

Figure 6.1: (a) Heart phantom (Model 067, CIRS Tissue Simulation Technology); (b) Volume rendering of the phantom from MR image.

of image effects that will not be included in my phantom data, such as the effects of breathing, cardiac motion and ECG gating errors.

6.1.1 Data acquisition on phantom

Four sets of 3D echo images were acquired from an Ultrasound Heart Phantom (Figure 6.1) using the same Philips iE33 ultrasound system with optical tracking as described in Section 4.4.1. My aim was to produce phantom datasets of the same object with varying image quality. The high quality images were used as a reference gold standard, which is compared to compounded volumes produced with lower quality (but much more clinically realistic) image data.

Dataset GS (gold standard) : My aim here was to obtain the best quality data possible from my phantom, simulating a situation where firstly there was no intervening tissue between the transducer and heart surface; secondly a very wide range of view angles were available (as seen in Figure 6.2 the heart phantom can be imaged from any

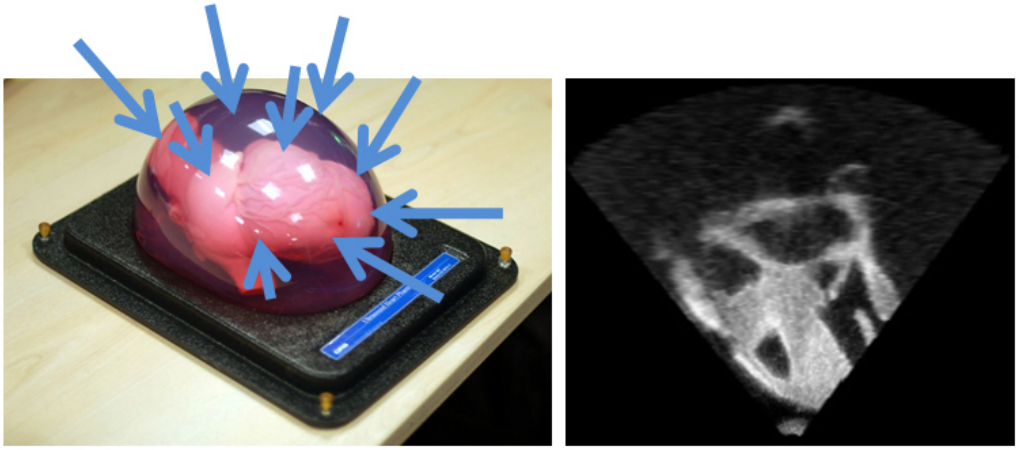


Figure 6.2: Illustration of acquiring a gold standard dataset from phantom. An typical echo image directly imaged from this heart phantom is shown on the right.

position on its upper surface); and thirdly there was no restriction on the total time used in data acquisition. Consequently 50 images were acquired from a wide range of angles with the transducer directly on the phantom surface. These images were used to produce a "best quality" compounded volume. They were also used as the reference images to calculate my image quality metrics as described in Section 6.1.2.

Dataset CV (clinical views) : This dataset is designed to simulate a more realistic clinical situation in terms of the echo view angles obtainable, and clinical time restrictions. Consequently only 10 images were acquired around apical 4 chamber view and parasternal long axis view (5 each as seen in Figure 6.3). Acquiring a number of images using small variations from standard clinical echo views has been proposed as a possible clinical protocol for image acquisition for compounding (Rajpoot et al., 2009).

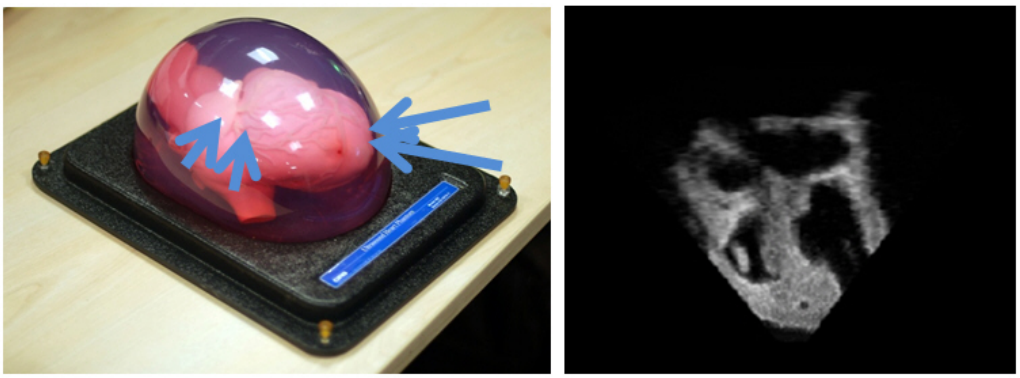


Figure 6.3: Illustration of acquiring a clinical relevant view dataset from phantom. An typical echo image is shown on the right.

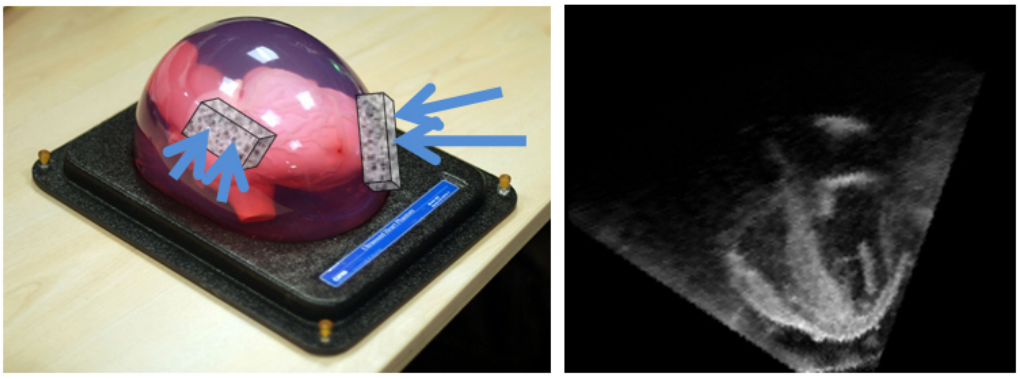


Figure 6.4: Illustration of acquiring a dataset with simulated soft tissue effect from phantom. A tissue mimicking phantom was used to degrade the image quality (shown on the right).

Dataset ST (soft tissue) : In this dataset clinical realism is further increased by simulating a situation where there is an attenuating layer of soft tissue between the transducer and heart surface. This was achieved using a tissue mimicking phantom, made from gelatine with talcum powder. It was placed between the heart phantom and transducer to simulate a layer of soft tissue, resulting in degraded image quality (Figure 6.4). The same clinically realistic view angles, and number of images (10) as Dataset CV were then acquired.

Dataset AT (artefacts) : In this dataset I additionally simulate artefacts that can come from the presence of bone (e.g. a rib) between

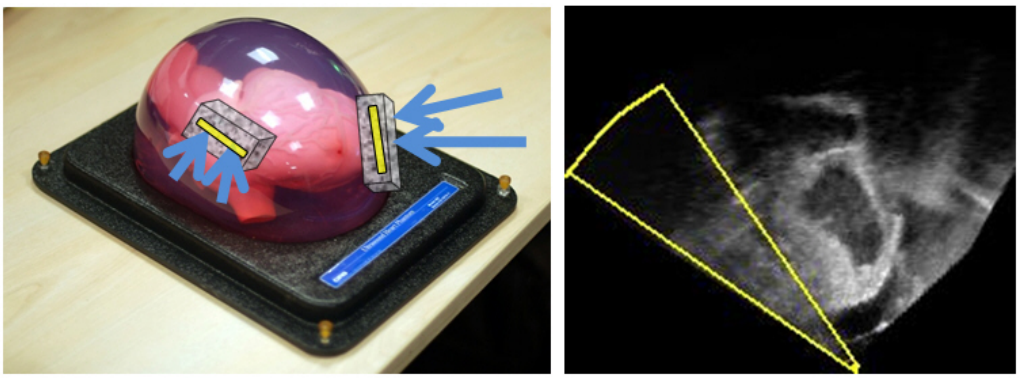


Figure 6.5: Illustration of acquiring a dataset with simulated artefact effect from phantom. An example of shadowing artefact which is caused by bone in tissue mimicking phantom (shown on the right).

the transducer and heart surface. This was achieved using a gelatine phantom as described above, but this time in addition it included a chicken bone. Again the gelatine with bone was placed between the heart phantom and transducer to degrade image quality and simulate artefacts (Figure 6.5). The same clinically realistic view angles, and number of images (10) as Dataset CV were then acquired.

6.1.2 Experiments on phantom data

My results use standard image quality indices, SNR and contrast. These indices are calculated using points and a region delineated in an original image as illustrated in Figure 6.6, where Label 1 is in homogeneous tissue regions (within myocardium), Label 2 is within septum tissue, close to the ventricle boundary and Label 3 is a region defined in the blood pool. An original image was chosen for this point picking process (rather than a compounded image) as to avoid this process biasing the result in favour of the chosen image.

I define the SNR as $SNR = \frac{\overline{Label1}}{\sigma_{Label3}}$, where $\overline{Label1}$ is the mean intensity

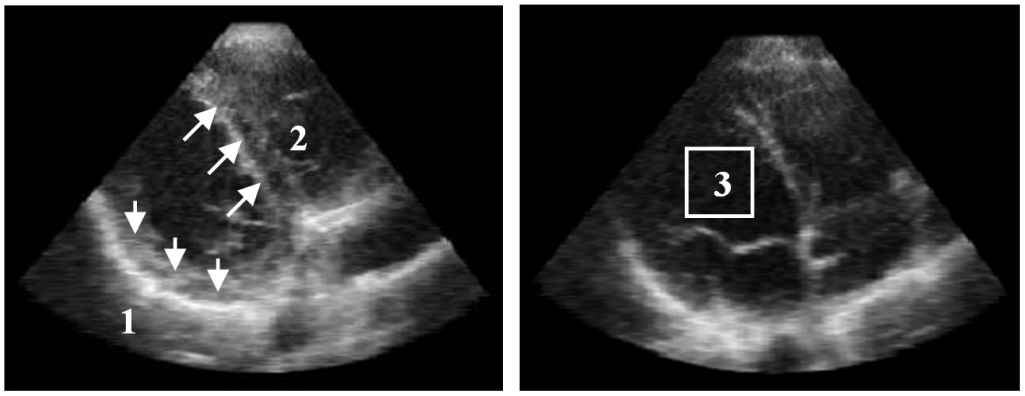


Figure 6.6: Overview of where points were picked to calculate SNR and contrast. Label 1, 30 points in homogeneous tissue regions (within myocardium) to obtain a mean signal intensity for calculating SNR; Label 2, 30 points picked within septum tissue, close to ventricle boundary to calculate contrast; Label 3, a region of size $10 \times 10 \times 10$ was defined in the blood pool to calculate standard deviation of image noise. Volunteer data V1 are presented.

value of the manually defined points (Label 1), and σ_{Label3} is the standard deviation of intensities on the region picked in the blood pool (Label 3). The contrast is defined as $contrast = \overline{Label2} - \overline{Label3}$, where $\overline{Label2}$ is the mean intensity value of the manually defined points within Label 2, and $\overline{Label3}$ is the mean value of intensities on the region picked in the blood pool. For the compounded volumes, the same points and regions are mapped into the compounded volumes using the registration transformation and SNR and contrast are calculated.

I report percentage changes in SNR and contrast, between my compounded images, and the original uncompounded gold-standard images. Therefore, a change in SNR and contrast for the phantom experiments of 0 represents ideal performance of the compounding algorithms, because this would mean it was able to overcome the effects of artefacts degrading to get back to the gold-standard image quality.

I compared my method to two basic statistical compounding methods: Mean and Maximum. These two methods were chosen because the Mean method is expected to perform well in terms of SNR by removing random noise, and the Maximum method should perform well in terms of contrast as discussed in Chapter 4. Therefore, Mean and Maximum were chosen as benchmarks. Experiments were carried out using all the phantom data. In addition, in order to further simulate clinical time restrictions, each of the datasets which contained 10 images were reduced down to just five images and the experiments re-run. Later in this section, my method is also compared to the best performing method (Phase Selection) from Chapter 4.

SNR and contrast show particular aspects of image quality. It is difficult to find a single measure which captures a range of these aspects. One such measure which has been proposed is a metric called the Universal Image Quality Index (UQI) (Wang and Bovik, 2002), which is designed to capture the overall image quality and similarity between a reference and a degraded image:

$$UQI_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \cdot \frac{2\bar{I}'_i \bar{I}'_j}{\bar{I}'^2_i + \bar{I}'^2_j} \cdot \frac{2\sigma_i \sigma_j}{\sigma_i^2 + \sigma_j^2} \tag{6.1}$$

where, the three components are cross correlation, mean intensities comparison and the variances between two images. The value $UQI \in [0, 1]$, and a high value denotes a high quality image. The UQI measure was calculated using phantom datasets (GS as the reference) between datasets GS&CV, GS&ST, GS&AT, firstly using the original individual images, and then for the feature consistency based compounded images. As the UQI measure requires a reference dataset it is not possible to calculate this measure on

the volunteer or patient datasets.

6.1.3 Results on phantom

Phantom dataset results are presented visually in Figure 6.7 and numerically in Figure 6.8. Figure 6.7 shows sample image slices from a single original image and slices from images compounded using each compounding method (Mean, Maximum, Proposed). Results from each dataset (GS, CV, ST and AT) are shown with 5 and 10 images. Both the Mean and Maximum methods performed as expected, blurring can clearly be seen in the Mean method images and high contrast noise in images produced using the Maximum method. My proposed method performed well, showing low noise and high contrast even as the datasets include more soft and hard tissue artefacts.

These visual observations are backed up by the numerical results in Figure 6.8. This figure shows the percentage improvement in SNR and contrast, compared to the uncompounded original gold-standard images. Not surprisingly, it shows decreases in image quality as datasets become worse. The Mean method achieved a high SNR improvement, but performed badly in terms of contrast. The opposite occurred with the Maximum method, which achieved good contrast improvement, but low SNR. The figure clearly shows the benefits of my approach, which has almost achieved the same SNR improvement as the Mean method while simultaneously achieving the same contrast improvement as the Maximum method. Interestingly the $dSNR$ and $dContrast$ (improvement on SNR and contrast between the compounded image and the original images) for my method us-

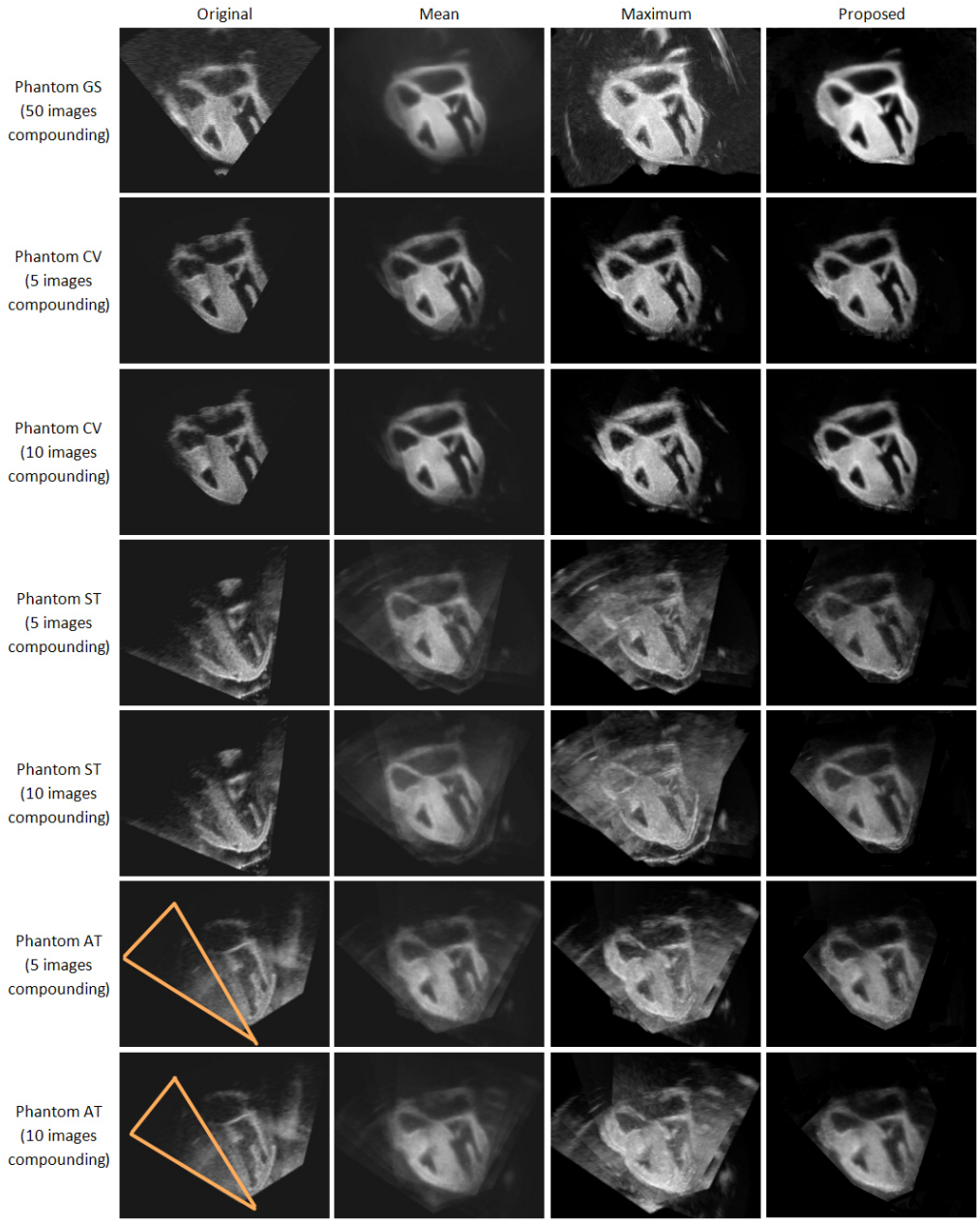


Figure 6.7: Slices from 3D echo images on phantom data. Original and compounded images produced using different numbers of images (5 images and 10 images) and different compounding methods: Mean, Maximum and my method (Proposed). My proposed method shows best combined results in terms of less noise (see Phantom ST), more contrast, and with artefacts removed (see Phantom AT labeled with triangle).

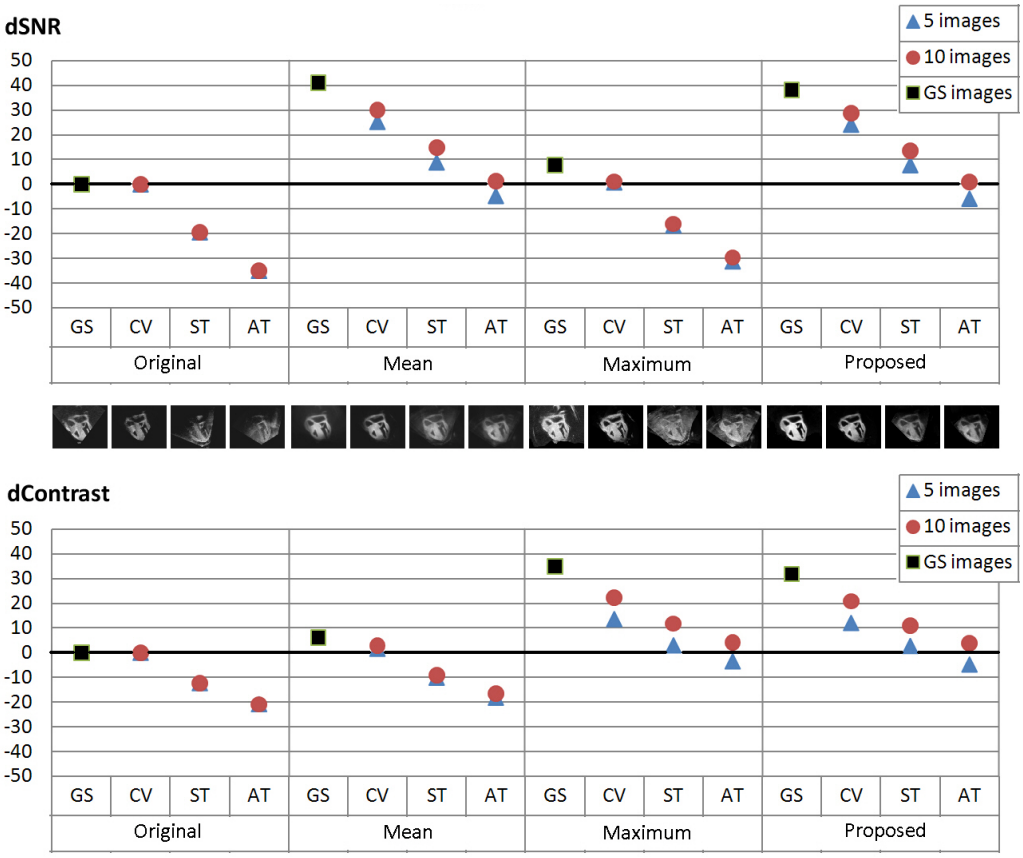


Figure 6.8: Phantom data, percentage changes in SNR and contrast between original images and compounded images produced using different numbers of images (5 images and 10 images) and different compounding methods: Mean, Maximum and my method (Proposed). All the values are calculated based on original gold-standard images.

ing the 10 image AT dataset are both approximately zero. Therefore, my compounding method has managed to take a set of ten images, which include a large amount of both hard and soft tissue artefacts, and produce a compounded volume with similar SNR and contrast to images acquired directly from the surface of the heart phantom.

In addition, my feature consistency method has been compared to the state-of-the-art phase-based methods, which has been extended to apply to multiple images (Phase Selection method described in Section 4.2.1). The results (in Figure 6.9 and numerically in Figure 6.10) show my feature consistency method has slightly better performance in terms of SNR and contrast.

Values of UQI between phantom original images, and between compounded images using my proposed method are shown in Figure 6.11. In each case the compounding 10 images have a higher UQI value compared to 5 images compounding and individual images, indicating that the compounded images have better overall image quality. In addition the percentage decrease in UQI measured image quality, as artefacts are introduced, is smaller for compounded using 10 images compared to 5 images compounding and original images. This indicates that my compounding has managed to reduce the effect of the introduced image degradation and artefacts.

6.1.4 Discussion on phantom validation

In this section, my feature consistency compounding algorithm has been evaluated using phantom data. By taking advantage of using the phantom to produce "gold-standard" images and compared to more clinical realistic

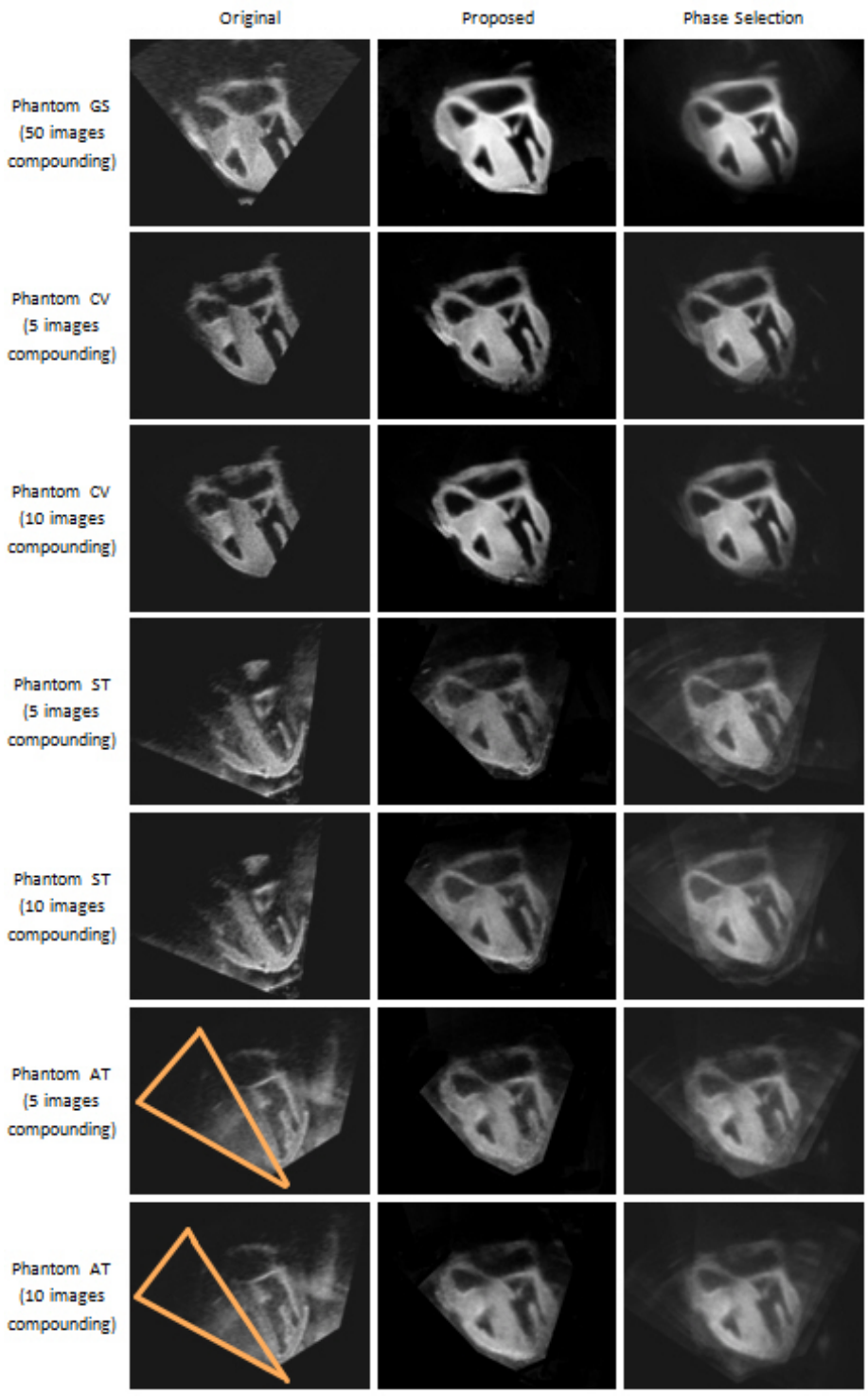


Figure 6.9: Slices from 3D echo images on phantom data. Original and compounded images produced using different numbers of images (5 images and 10 images). My method (Proposed) has been compared to the phase based compounding. My proposed method shows slightly better combined results in terms of less noise (see Phantom ST), more contrast, and with artefacts removed (see Phantom AT).

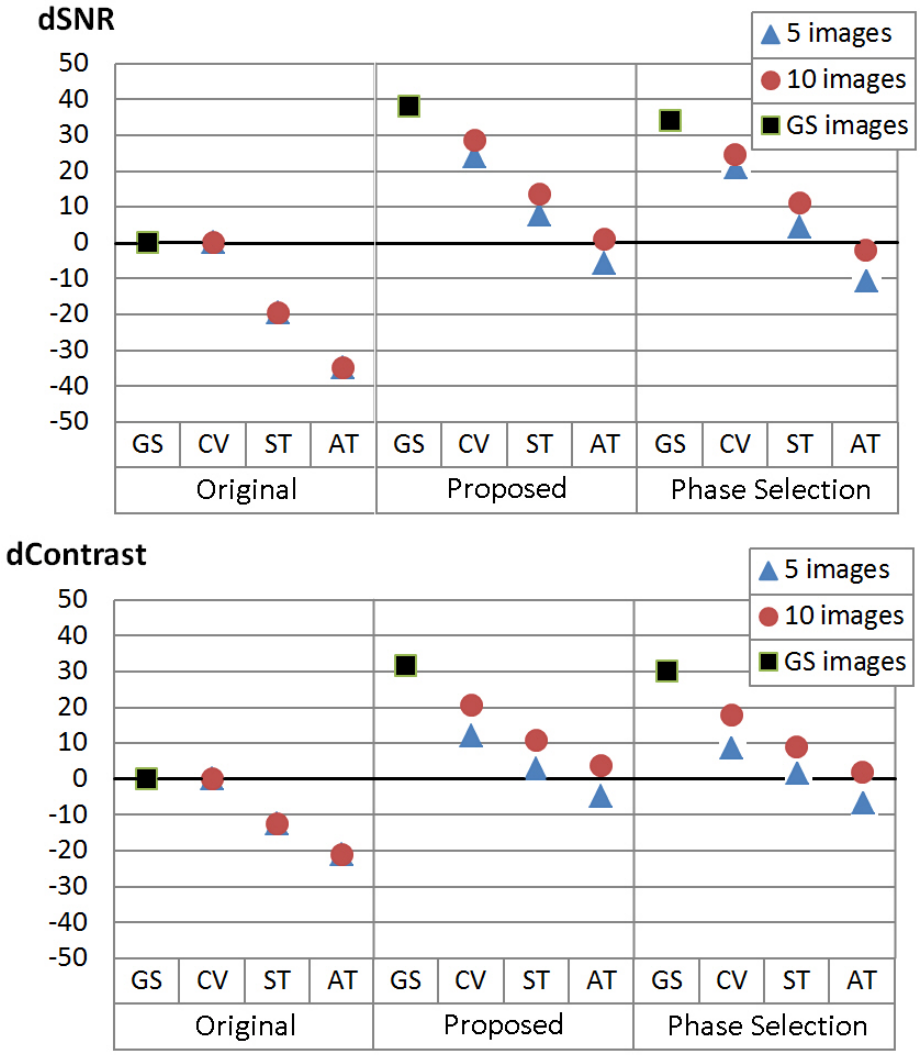


Figure 6.10: Phantom data, percentage changes in SNR and contrast between original images and compounded images produced using different numbers of images (5 images and 10 images). My method (Proposed) has been compared to the phase based compounding. All the values are calculated based on original gold-standard images.

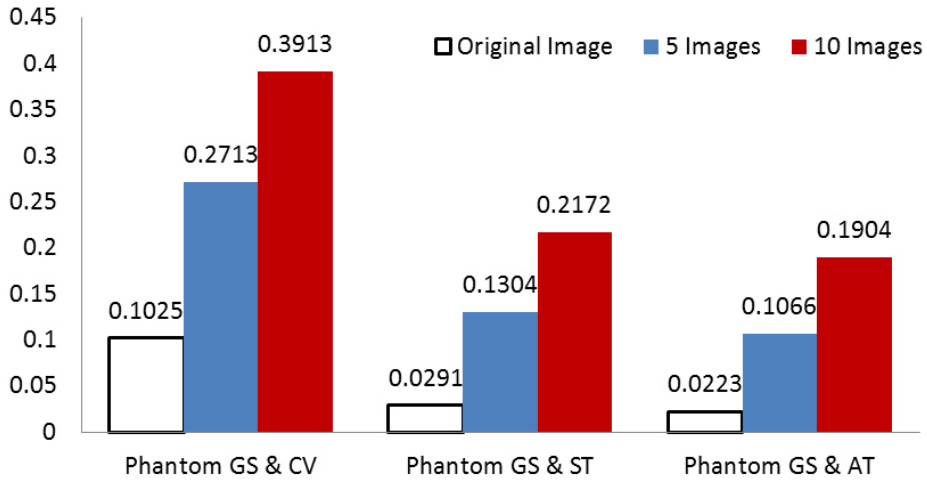


Figure 6.11: Phantom data, percentage changes in UQI between original images and compounded images produced using different numbers of images (5 images and 10 images). My method (Proposed) has been used.

phantom images, the phantom results show my method is able to take a set of ten images, degraded by soft and hard tissue artefacts, and produce a much higher quality compounded volume (equivalent to imaging directly on the phantom surface). The compounded image has reduced echo image artefacts, as well as decreased noise and improved boundary definition. Compared to Mean, Maximum and the previous Phase Selection method used in Chapter 4, my proposed method shows better quality in terms of SNR and contrast improvement.

6.2 Validation using volunteer data

In this section I will describe the procedure of using volunteer data to evaluate my feature consistency based compounding method. Compared with validation using phantom data, the volunteer echo images contains real cardiac motion and artefacts. However, the acoustic window is limited and it is not possible to acquire an "gold-standard" from volunteers.

6.2.1 Data acquisition on volunteer

The volunteer data was acquired with the volunteers in the left lateral decubitus position. A starting position for the registration was again obtained by optical tracking the transducer. For better acoustic access the echo cardiographer would often ask the volunteers to hold their breath at different respiratory positions. On a few rare occasions (which were noted at the time of acquisition) the volunteer was asked to roll either more onto their back or side. This movement was requested to be kept to a minimum as the volunteer is not tracked and it could effect the performance of the registration algorithm (move the starting estimate outside capture range).

The iE33 ultrasound system was again used, and the clinicians were asked not to adjust the time-gain compensation setting between acquisitions. ECG gating is routine during echo acquisitions and this was used to label each image within the cardiac cycle. Triggered by ECG gating, images were acquired over 4-7 cardiac cycles to produce wide field-of-view 3D images from the apical four chamber view and parasternal long axis view. Images at the end-diastolic phase were used for image registration and compounding. 10 datasets (V1-V10) were acquired from volunteers with tracking and probe calibration, with 10 images in each dataset. Image sizes varied depending on depth settings, typical spatial dimensions were $224 \times 208 \times 201$ voxels with a voxel size of $0.83 \times 0.82 \times 0.75\ mm^3$.

6.2.2 Experiments on volunteer data

Results were calculated using the same image quality indices, SNR and contrast, and on the same selected anatomic features as described in Section

6.1.2. For the volunteer datasets, because no gold-standard is available, I report a change in SNR and contrast between the compounded image and the original images in each dataset. Therefore, in these cases a positive SNR and contrast improvement represent good performance. In addition, as no "gold-standard" was available for volunteer data, I could not use the UQI measure.

In addition to the standard quality indices described above I also carried out a visual assessment validation on the volunteer data. The standard quality indices are useful to obtain quantitative measures of image quality and, as they are standard measures, can allow comparison between techniques. However, there are a number of other important measures of image "quality", some of which are difficult to assess quantitatively. In this visual assessment experiment the four criteria given in Table 6.1 were used to assess image quality. These criteria were proposed by our clinical colleagues and have all been previously used by other researchers to assess echo image quality.

From the literature, previous subjective echo visual assessment experiments normally use category scaling (Fetics et al., 2001; Mrak et al., 2003; Walsh et al., 2005; Mitra et al., 2006; Szmigielski et al., 2010). These experiments assign image quality ranked values to each image individually, e.g. "good", "intermediate" or "poor". However, because the aim of my compounding algorithm is to improve image quality, I have used an assessment experiment where the observer ranks images against each other rather than assigning a score to each individual image. I have chosen this methodology for a number of reasons:

Table 6.1: Visual assessment criteria for echo image quality.

Name	Description	Previously used by
Overall Image Quality (OIQ)	A subjective judgement based on the first instant observation about the overall image presentation, which considers all aspects of the image together, e.g. contrast, noise, anatomical feature completeness, etc.	Mulvagh *et al.*, 1998
Useable Field-of-View (UFOV)	The coverage of the valuable anatomical feature/information presented in a image.	Rajpoot *et al.*, 2011; Noble *et al.*, 2011
Endocardial Border Definition (EBD)	A good EBD will show good contrast between the blood pool and endocardial border in all segments of the endocardium. A poor EBD normally suggests poor contrast or missing endocardial boundaries.	Mulvagh *et al.*, 1998; Fetics *et al.*, 2001; Szmigielski *et al.*, 2010; Rajpoot *et al.*, 2011; Noble *et al.*, 2011
Cavity Noise (CN)	A low cavity noise image will show the blood pool as a uniform black region. A high cavity noise image will contain multiple white/grey speckles in the blood pool causing reduced contrast between the blood pool and endocardium.	Belohlavek *et al.*, 1998; Rajpoot *et al.*, 2011

1. My algorithm is designed to accept a number of original images, and then produce a compounded image of improved quality. Therefore, the important aspect I wish to measure is quality improvement between original and compounded images. My validation method is specifically designed to show this quality improvement by allowing the observers to rank all the images (original input images and compounded images) in order of quality.

2. Different observers can have different views on what constitutes a "good", "intermediate" or "poor" image, which belongs to a more general issue called score normalisation (Poh, 2011; Kenna and Berche, 2011; Rua et al., 2008). Thus adding additional confounding errors into my statistical analysis. Whereas there is likely to be less variation in opinion between observers as to whether one image is of better quality than another.

3. It is important to have expert observers carry out the visual assess-

ment validation. I have designed my experiment to make best use of their valuable time, allowing the assessment of large number of images by a reasonable (three) number of observers. Assessment schemes which required all the images to be individually assessed would been more time consuming resulting in fewer images assessed.

Therefore, for my visual study I proposed a revised method. Three clinicians experienced in echo were invited and each dataset was assessed separately. For each dataset two compounded volumes were produced and seven of the original images were selected. The compounded volumes were produced using 10 images (defined as image group "CMP_10") and using 5 images (image group "CMP_5"). The seven original images (image group "ORIG") were chosen by discarding the lowest quality images (determined by the most experienced clinician by visual inspection). A 2D plane (such as standard apical or parasternal view) was picked from one of the original 3D images in each image group by the most experienced clinician. All the images in that dataset were then transformed into the same coordinate system and corresponding 2D slices were extracted. A computer-based slide show was produced, which contained a separate slide for each dataset. Each slide contained the nine 2D slices for each single dataset, arranged in a random order. Three clinicians were independently invited to pick the best two images from each dataset ranked in order, according to the four criteria OIQ, UFOV, EBD and CN. Therefore the total number of images visually assessed for the volunteer datasets is 630 (10 dataset $\times$ 9 images presented per datasets $\times$ 4 quality criteria $\times$ 3 observers).

It is possible that a number of observer biases could have occurred due

to my chosen methodology. In a visual inspection validation methodology it is very important to use expert observers to assess the image quality. Such experts will be able to identify the differences between original and compounded echo images and so it is impossible to totally blind the observers during the validation process. A number of design choices are made when constructing the validation methodology which may also introduce observer bias. It is therefore extremely difficult to remove all possible sources of bias from such a visual clinical evaluation process. The steps I have taken to try to minimize observer bias are discussed below:

1. Bias in the choice of reference slice. By requiring that the chosen slice was based on one of the original, standard view images, rather than allowing it to be chosen from the compounding image, it is hoped that if bias is introduced here it will be in favour of the original images. Therefore, bias introduced should not be able to unfairly improve the results from the compounding algorithm.

2. Bias because a single clinician chose the reference slice. The most experienced observer was responsible for choosing the 2D slice plane to use. This should have resulted in the slice with the most clinically relevant information being chosen. Therefore bias caused by the choice of slice should be towards compounding enhancing the most relevant clinical information which is a good feature of the chosen methodology.

3. Bias because the observers can see which images are compounded. As has been detailed earlier, my compounded algorithm aims to improve image quality and FOV. A knowledgeable observer (e.g. an echocardi-

ographer) of ultrasound images will be able to detect these differences and will know which are the compounded and original images. Even if the FOV was normalised the compounded images could be detected by their different texture. Although I have tried to randomise the position of original and compounded images, and to include an observer who was familiar with compounded images and observers who had never seen images compounded by my algorithm before, this bias still cannot be avoided. It is difficult to assess which way this bias will act. However if the bias occurs it is most probable that the clinician most familiar with compounded images will give different ranked scores compared to the other observers. In order to assess this I carried out the Friedman test (Friedman, 1937), a nonparametric procedure to compare more than two paired samples. This was used to find out whether the clinician most familiar with compounded images performs significantly differently from the other two clinicians, for each quality criteria described before.

The work described in this paper is focused on spatial compounding, and so I have compounded end-diastolic images, as the heart is moving slowly, to reduce the chance of temporal misalignment effecting compounded image quality. However, I have used my algorithm to compound a number of 4D volumes. Temporal alignment was achieved as follows (Figure 6.12): All my acquisitions were ECG gated. The number of volumes acquired at each view direction depended on the current heart rate, a faster heart rate resulting in fewer volumes. Each image was labelled with its fractional position within the cardiac cycle, for example if ten images were acquired

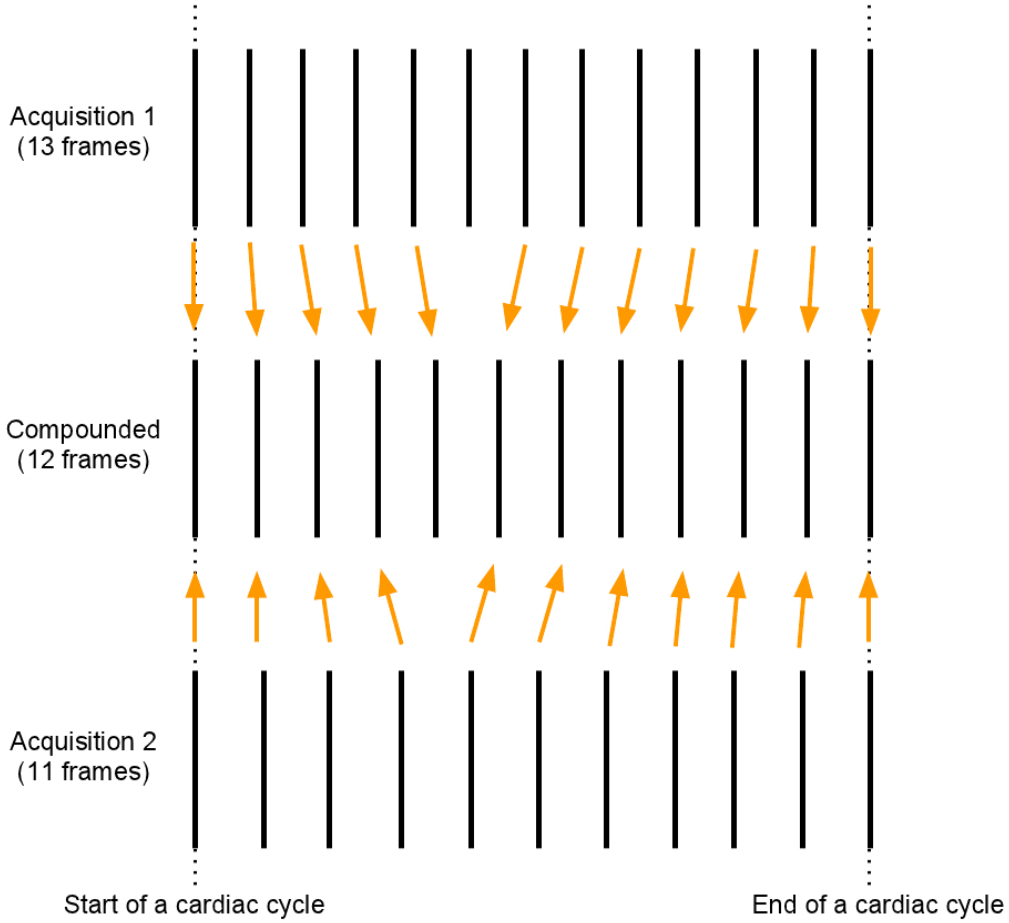


Figure 6.12: Temporal alignment of 4D volumes. Take two acquisitions for example. Acquisition 1 has 13 frames in one cardiac cycle, and acquisition 2 has 11 frames. Therefore, in the compounded data should contain the median value of the frames, which is 12. Then frames from both of the acquisitions will be aligned using nearest neighbour interpolation (as arrows indicated).

from a given view direction the images would be labelled as 1/10, 2/10, . . . , 10/10. The number of frames in the compounded 4D volume was set to the median number of frames in the input data sets. Temporal correspondence between sets of images from different view direction was achieved using nearest neighbour interpolation based on the fractional position of each image within the cardiac cycle.

6.2.3 Results on volunteer data

Volunteer datasets results are presented visually in Figure 6.13 and numerically in Figure 6.14, and are consistent with the phantom results. To demonstrate the performance of my proposed method, example images have been chosen from a volunteer dataset in Figure 6.13. To highlight differences in performance, regions of interest have been chosen (yellow box) and magnified. These were chosen at the endocardial border in volunteer dataset V4. Again, my method shows a better performance compared to the other compounding methods in terms of achieving better myocardium definition, less cavity noise and good contrast. The numerical SNR and contrast improvement values in Figure 6.14 show significant differences when compounding using 10 images compared to 5 images ($p - value < 0.0001$ using my proposed method).

Figure 6.15 shows the results for the visual assessment experiments. The results are presented as the percentage of times each image group ("CMP_10", "CMP_5", "ORIG") were ranked first (Figure 6.15a) or second (Figure 6.15b) out of the set of nine images presented to each clinician. The bars are split into the four categories which were investigated as described in section 6.2.2. The compounded volumes using 10 images have clearly been ranked as being the best for each category ($p - value < 0.001$). The second ranked image group is then dominated by the compounded volumes using 5 images. For the interobserver variability, the Friedman test has show $p - value > 0.1$ (for each quality criteria), which means there is no statistically significant difference and the observations are considered to be consistent for all the three clinicians.

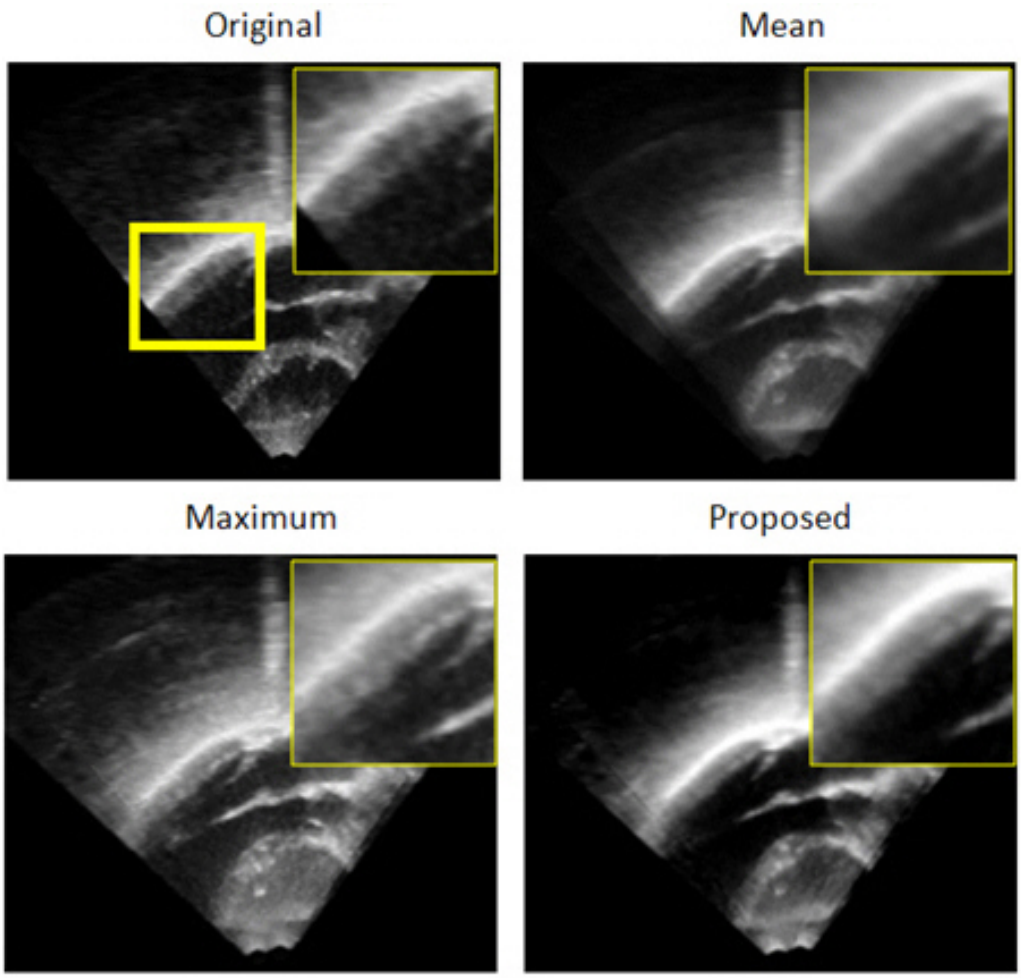


Figure 6.13: Slices from single original images, and compounded images using maximum images available and different compounding methods: Mean, Maximum and my method (Proposed). Volunteer dataset V4 has been shown. The specific feature highlighted is the Endocardial Border.

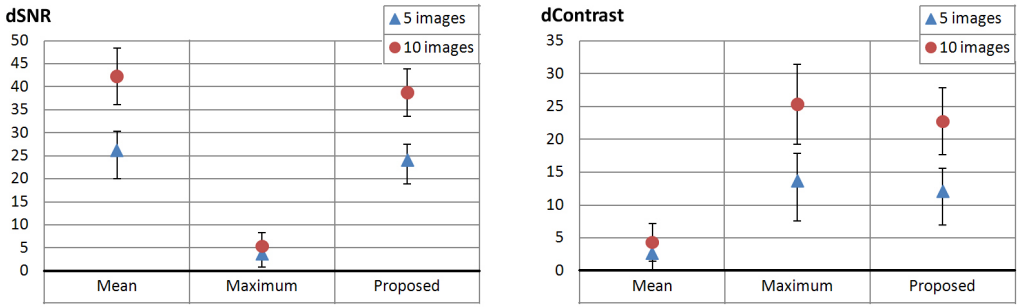
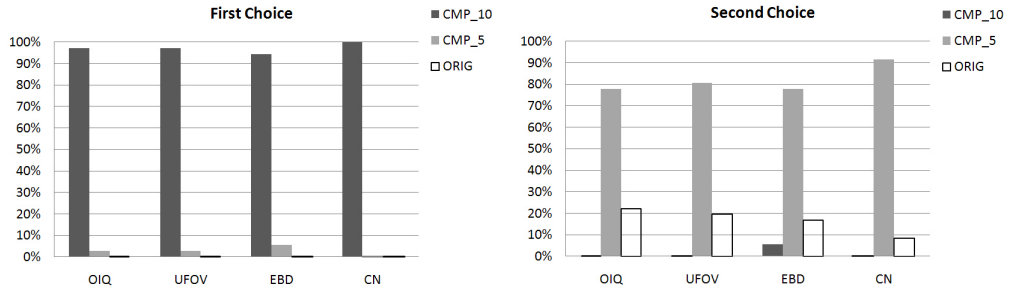


Figure 6.14: Volunteer data, percentage changes in SNR and contrast between original images and compounded images produced using different numbers of images (5 images and 10 images) and different compounding methods: Mean, Maximum and my method (Proposed).



(a) (b)

Figure 6.15: Clinical visual inspection results. Percentage of times an image was ranked the best (a) or second best (b) from nine presented images for each dataset. Presented images were the compounded volume with ten images CMP_10, compounded with five images CMP_5, and the seven best original images ORIG. Categories for selection were overall image quality (OIQ), useable field of view (UFOV), endocardial border definition (EBD) and cavity noise (CN).

More results on the effect of using more images are presented in Figure 6.16 which gives a visual example of the improvement in image quality using an increased number of images. The figure shows how the field-of-view is increased, noise is reduced, definition of septum and endocardial border is enhanced and coherent homogeneous tissue regions are produced. Figure 6.17 shows numerically how compounding using an increased number of images improves SNR and contrast, again indicating the advantage of using a greater number of images.

I also used my algorithm to compound a number of 4D volumes. A 2D slice from eight temporal frames of one such volume are shown in Figure 6.18. The image sequence shows good contrast, particularly between the myocardium and the blood pool, which is essential for the accurate assessment of the endocardial border.

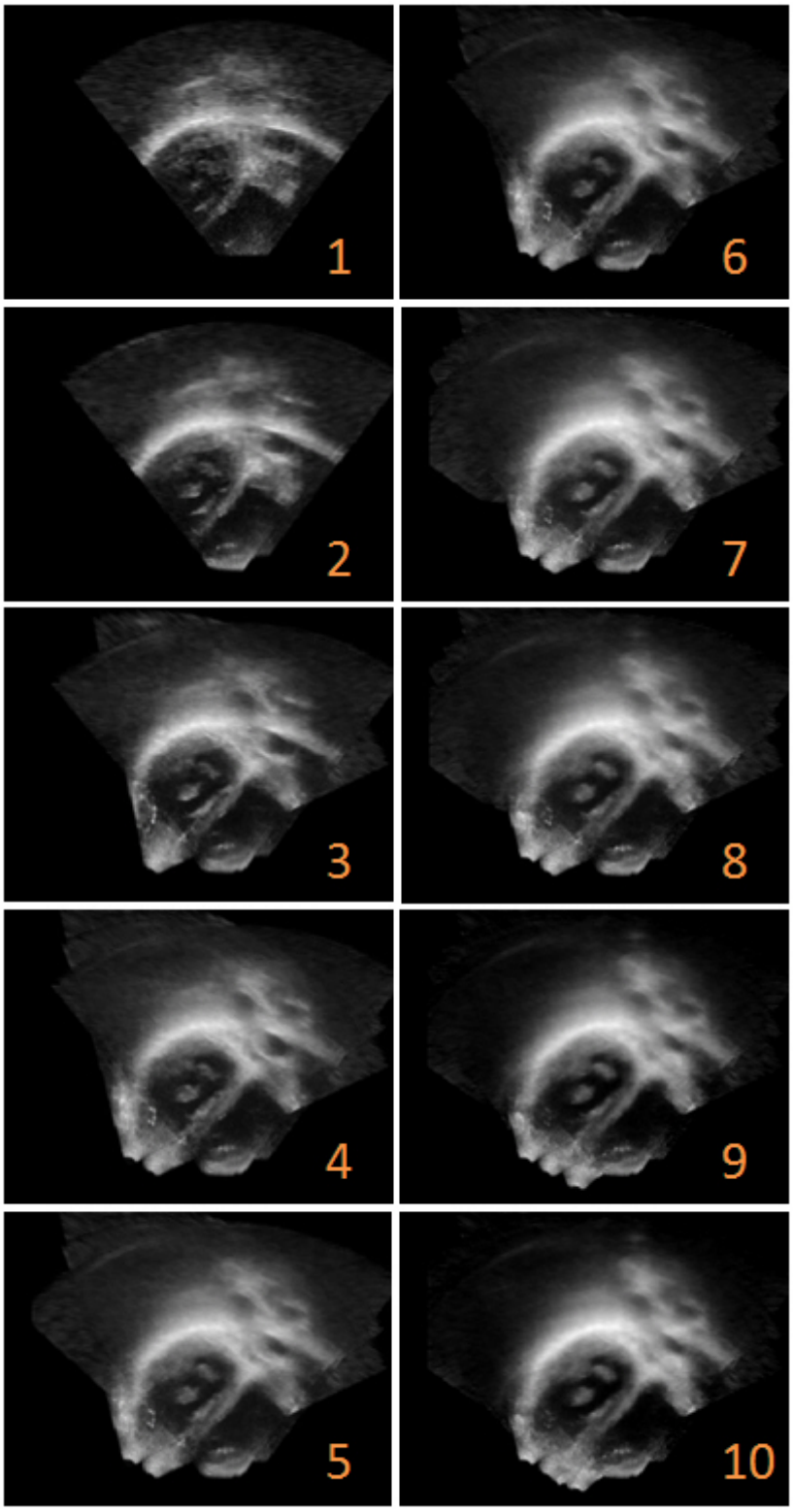


Figure 6.16: Visual example of how compounding larger numbers of images can improve quality. Top left shows an original single image, and then compounded volumes are produced using an increased number of images with maximum of 10 images. Volunteer data V6 is demonstrated. My proposed compounding method is used to show the improvement on the definition of septum and endocardial border, and how field-of-view is extended.

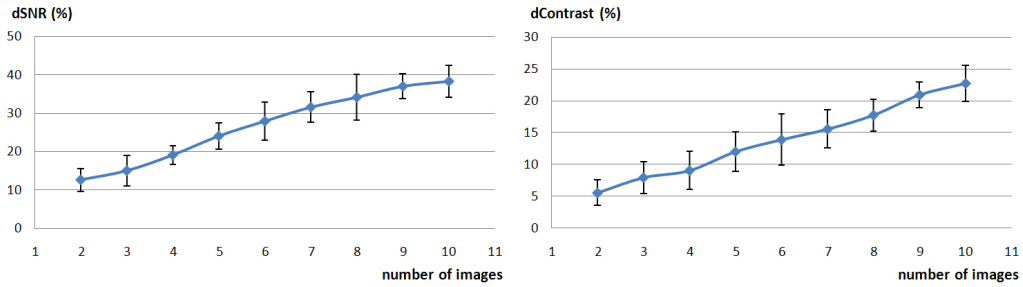


Figure 6.17: Effect of compounding using an increased number of images on image quality. Graphs show mean percentage change in SNR (left) and contrast (right) calculated over all volunteer datasets, the error bars show standard deviation. My proposed compounding method is used.

6.2.4 Discussion on volunteer validation

The volunteer data has been used for evaluating my consistency compounding algorithm. Numerical indices have been used to measure image quality, and subjective visual inspection is carried out by experienced clinicians to rank image quality. The visual inspection studies by three clinicians showed a very strong preference for my compounded volumes in terms of overall high image quality, large field-of-view, high endocardial border definition and low cavity noise.

My method to calculate temporal correspondence (nearest neighbour interpolation based on position in cardiac cycle) produced visually realistic 4D compounded volumes (Figure 6.18). However, more sophisticated methods to obtain temporal correspondence would probably greatly improve 4D image quality. In particular, full 4D registrations could be carried out, allowing non-linear alignment in the temporal dimension to account for variations in cardiac cycle length. Similar approaches have been used (though inter-subject) to allow full 4D spatio-temporal free form registration of cardiac MR images (Perperidis et al., 2005), and for 4D myocardial

Figure 6.18: Example slices from one cardiac cycle of a 4D compounded image sequence on Volunteer dataset V6. These images show my proposed 3D compounding technique has been applied to each cardiac phase to produce a 4D compounded sequence, and they seen to contain lower noise, fewer artefacts, more constant intensities within tissue homogenous region, and sharper boundary definition.

motion and strain estimation (Piella et al., 2011). The ability to account for variations in cardiac motion will likely become more important when using patient data with arrhythmia, or to compound fast moving structures such as valves.

6.3 Overall Discussion

In this chapter, the validation method has used phantom and volunteer datasets. Results from both phantom and volunteer experiments show that my feature consistency based compounding algorithm improves image quality. Results also show that the use of ten as opposed to five images is significant with respect to numerical measures of image quality and from visual inspection. There is obviously be a limit to the number of volumes that can be acquired as a routine clinical practice. One method which could allow the acquisition of more images within a clinically acceptable time frame would be to use real time 3D data streaming. This would produce many more images, but they would each have smaller field of view and would likely be of lower quality than the images acquired and used in this submission. This would allow the compounding algorithm to use much more data, but would require the algorithm to have greater ability to register accurately with noisy small field-of-view data.

In the next chapter, an initial investigation will be carried out to ascertain whether compounding is able to provide clinical benefit. Potential areas of benefit include the extended field-of-view which may permit better and more complete imaging of some parts of the heart, which are difficult or

impossible to incorporate into a single echocardiographic projection, such as the right ventricle.

Chapter 7

Clinical applications

In previous chapters, my phantom and volunteer experiments have shown how my compounding technique improves 3D echo image quality. This chapter is my initial step to assess how my compounding technique performs using clinical data.

When moving from volunteer to patient data there are 3 main issues involved with respect to compounding algorithm performance:

- Firstly my method to obtain starting estimates for my registration. Our optical tracking equipment which has been used in my phantom and volunteer experiments is not a standard addition to an echo machine. This tracking device is not practical to be deployed in clinical routine in my institution during the time frame of my PhD, due to time and space constraints in clinic and ethical considerations. In addition, even if a transducer is tracked, patient movements, which are probably more likely to occur with patients than volunteers, will cause errors in my starting estimates.

- Secondly there will likely be increased problems in obtaining temporal

alignment between my input images in a dataset. Although I concentrate on spatial alignment between different view images, patients who have an irregular heart rhythm will be likely to cause increased temporal alignment errors.

- Thirdly elder adult patients normally have worse acoustic windows compared to healthy young volunteers. Therefore poor quality or limited field-of-view images will be acquired, which can make image registration difficult.

In this chapter, I will concentrate on two types of pathology: Transposition of the Great Arteries (TGA) and Hypoplastic Left Heart Syndrome (HLHS). The reason for choosing TGA is that it is a regularly observed pediatric case, which is similar to some other pathologies and is therefore a good example to use for testing my compounding algorithm. HLHS is chosen as a potential clinical pathology that could benefit from compounding, as compounding multi-view images will provide wider coverage and higher quality image. This is particular useful for the right ventricle, which could allow better function assessment of HLHS patients.

My consistency based compounding strategy will be applied to TGA and HLHS cases, to see how my compounding algorithm works on these clinical data. In the rest of this chapter, separate experiments will be carried out, firstly to compound TGA datasets, and then HLHS datasets. The chapter ends with an overall discussion on the use of compounding on these specific datasets and on clinical data in general.

7.1 Transposition of the Great Arteries (TGA)

7.1.1 Clinical Background

A normal heart should have the great arteries connected in the following way (see Figure 7.1 for illustration): the pulmonary artery connected to the right ventricle which carries deoxygenated (blue) blood from the right ventricle to the lungs; the aorta is connected to the left ventricle which carries oxygenated (red) blood from the left ventricle to the body. However, a patient with transposition of the great arteries (Martins and Castela, 2008), has the pulmonary artery and aorta exchanged their connections. In normal practice, clinicians diagnose TGA by looking at the heart from a number of different probe positions. Different anatomical structures are visualised in each of the different probe positions. No single dataset is generated which incorporates all diagnostic information. An advantage of the compounded multi-view images is that images taken from different probe positions could be compounded into a single data which could be interrogated to produce all necessary views. Furthermore, this dataset could be retrospective reinterrogated to produce novel images which were not obtained at the time of the initial clinical study.

7.1.2 Experiments

Two previously acquired datasets (TGA1-2) from TGA patients with 5 images per dataset were used. Typical volumes were $256 \times 272 \times 201$ voxels with a $0.38 \times 0.38 \times 0.33\ mm^3$ voxel size.

My feature consistency compounding method (proposed in Chapter 5)

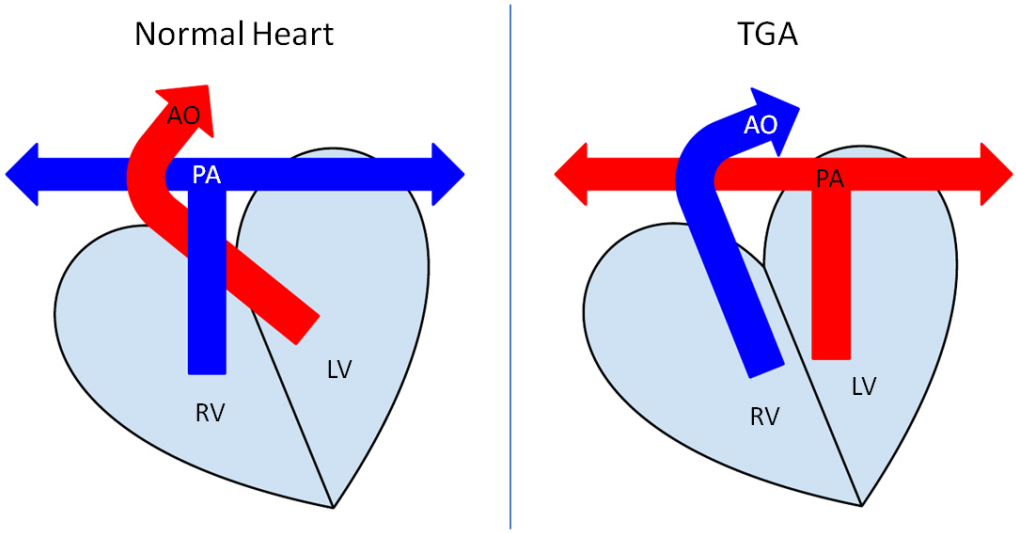


Figure 7.1: Simplified illustration of the difference between a normal heart and a patient heart with TGA. (LV = left ventricle; RV = right ventricle; PA = pulmonary artery; AO = aorta)

was applied to the TGA patient data. As the transducer was not tracked, instead three landmarks (on valves and apex) were manually picked by an experienced clinician. These were used as input to a Procrustes point-based registration algorithm implemented using the Visualization Toolkit (VTK, Kitware Inc., NY, USA) to calculate a starting estimate for my intensity-based registration algorithm RCF which was described in Section 5.4.1.

For validation, a visual assessment of compounded image quality was carried out (as described in Section 6.2.2). Three clinicians are invited to pick the best two images from a selection which included each of the input images and a compounded image. The same criteria was used as before: Overall Image Quality (OIQ), Useable Field-of-View (UFOV), Endocardial Border Definition (EBD), and Cavity Noise (CN).

Similar to Section 6.2.2, I have used my algorithm to compound 4D volumes.

7.1.3 Results

Example images are presented in Figure 7.2 and Figure 7.3. These images are orientated into the standard view and the TGA specific view angles, which are 4CH (apical four chamber) and LV2PA (left ventricle to pulmonary artery). 4CH views in both figures show the overview of the heart. The compounded images have clearer endocardial border definition (labelled in Figure 7.2 as EB) and wider field-of-view coverage, which are very good for function assessment. In particular, the LPA and RPA features are very well presented at LV2PA view angle in Figure 7.2, which are typically very difficult to see and acquire in a single echo volume. All my compounded volumes have been presented with improved SNR and contrast, compared to individual input images.

Registration was seen to be successful for all the images using visual inspection. The clinical image quality visual assessment study results are shown in Figure 7.4. These show very similar results to my volunteer study where again the clinicians showed a strong preference for the compounded image rather than individual input images. The compounded image was always chosen as the best image over all four criteria: OIQ, UFOV, EBD and CN.

I also used my algorithm to compound 4D volumes of these datasets. A 2D slice from eight temporal frames of one such volume is shown in Figure 7.5. The image sequence shows good contrast, particularly between the myocardium and the blood pool, which is essential for the accurate assessment of the endocardial border and the function of the heart. Current temporal compounding is very basic (refer to Section 6.2.2). However, it

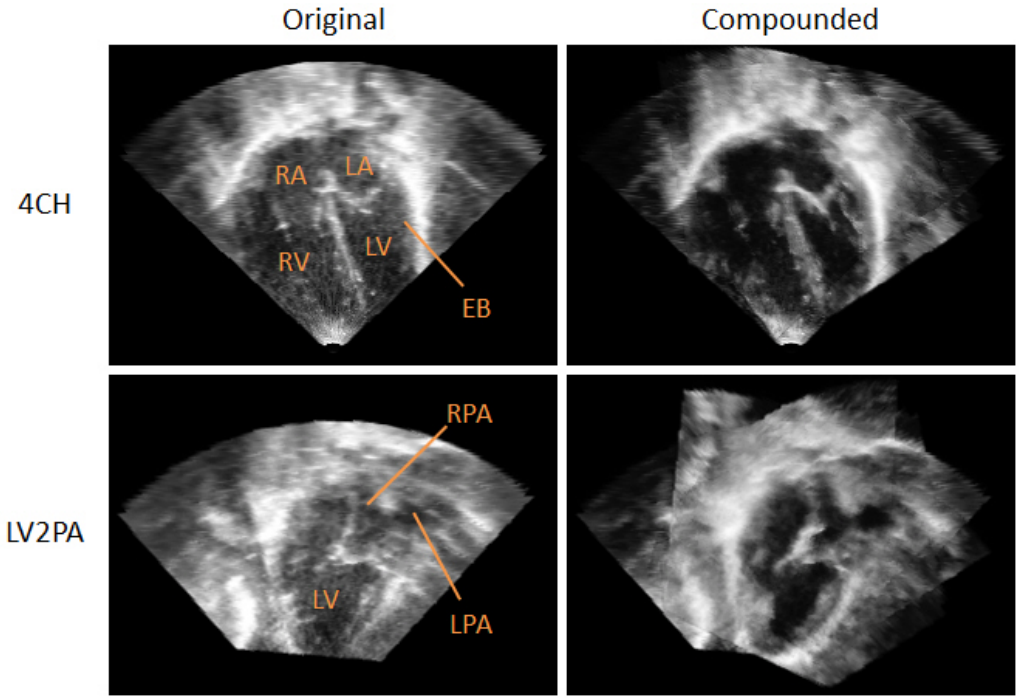


Figure 7.2: Slices from 3D echo images on TGA1 patient data. Original and compounded images are presented in the following view angles: 4CH (apical four chamber) and LV2PA (left ventricle to pulmonary artery). My compounded images shows less noise, more contrast and with all TGA features well presented in single volume. (LV = left ventricle; LA = left atrium; RV = right ventricle; RA = right atrium; EB = endocardial border; LPA = left pulmonary artery; RPA = right pulmonary artery)

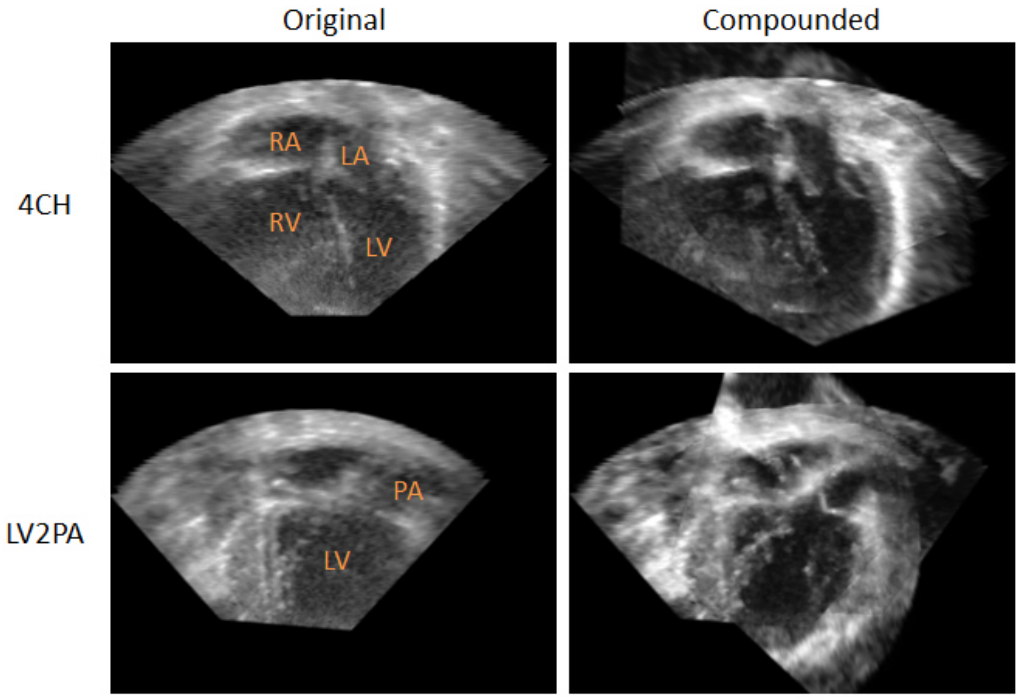


Figure 7.3: Slices from 3D echo images on TGA2 patient data. Original and compounded images are presented in the following view angles: 4CH (apical four chamber) and LV2PA (left ventricle to pulmonary artery). My compounded images shows less noise, more contrast and with TGA features well presented in single volume. (LV = left ventricle; LA = left atrium; RV = right ventricle; RA = right atrium; PA = pulmonary artery)

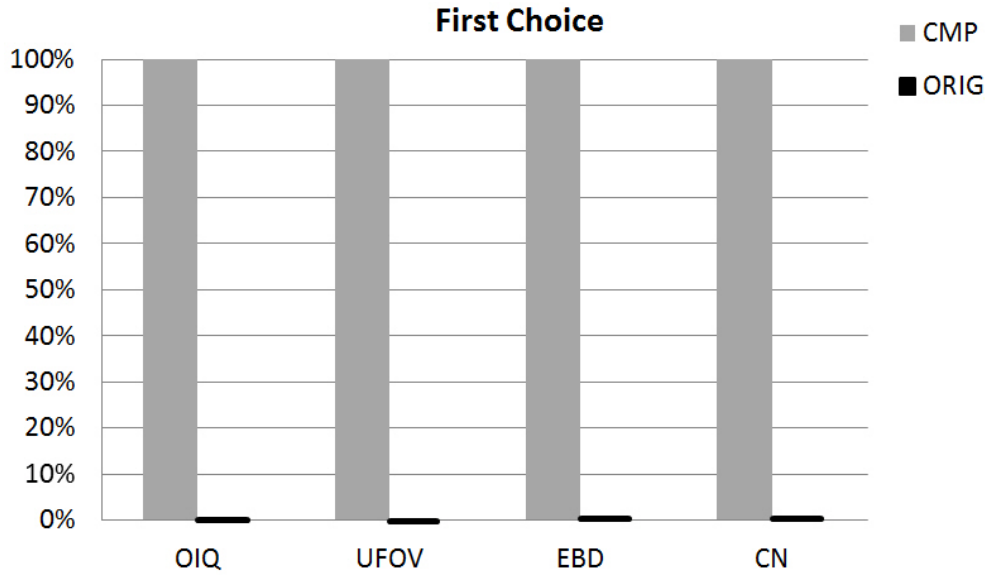


Figure 7.4: Clinical visual inspection results. Percentage of times an image was ranked the best from presented images for each dataset. Presented images were the compounded volume CMP and the original images ORIG. Categories for selection were overall image quality (OIQ), usable field-of-view (UFOV), endocardial border definition (EBD) and cavity noise (CN).

still produced a visually plausible 4D movie.

7.1.4 Discussion

In this section, two TGA datasets have been processed by my 3D echo compounding techniques. Moving to clinical data, my compounding algorithm has continued to perform well, showing a similar good performance as with the volunteer data (in Section 6.2.3).

Although no tracking information has been provided for image alignment, the manually selected landmarks on images still provided an accurate enough starting estimate for my registration algorithm. However, point picking could not be done routinely as it needs time and manual input from a clinician. TGA patients do not usually have an irregular heart rate. Therefore, using a basic temporal alignment technique will still produce

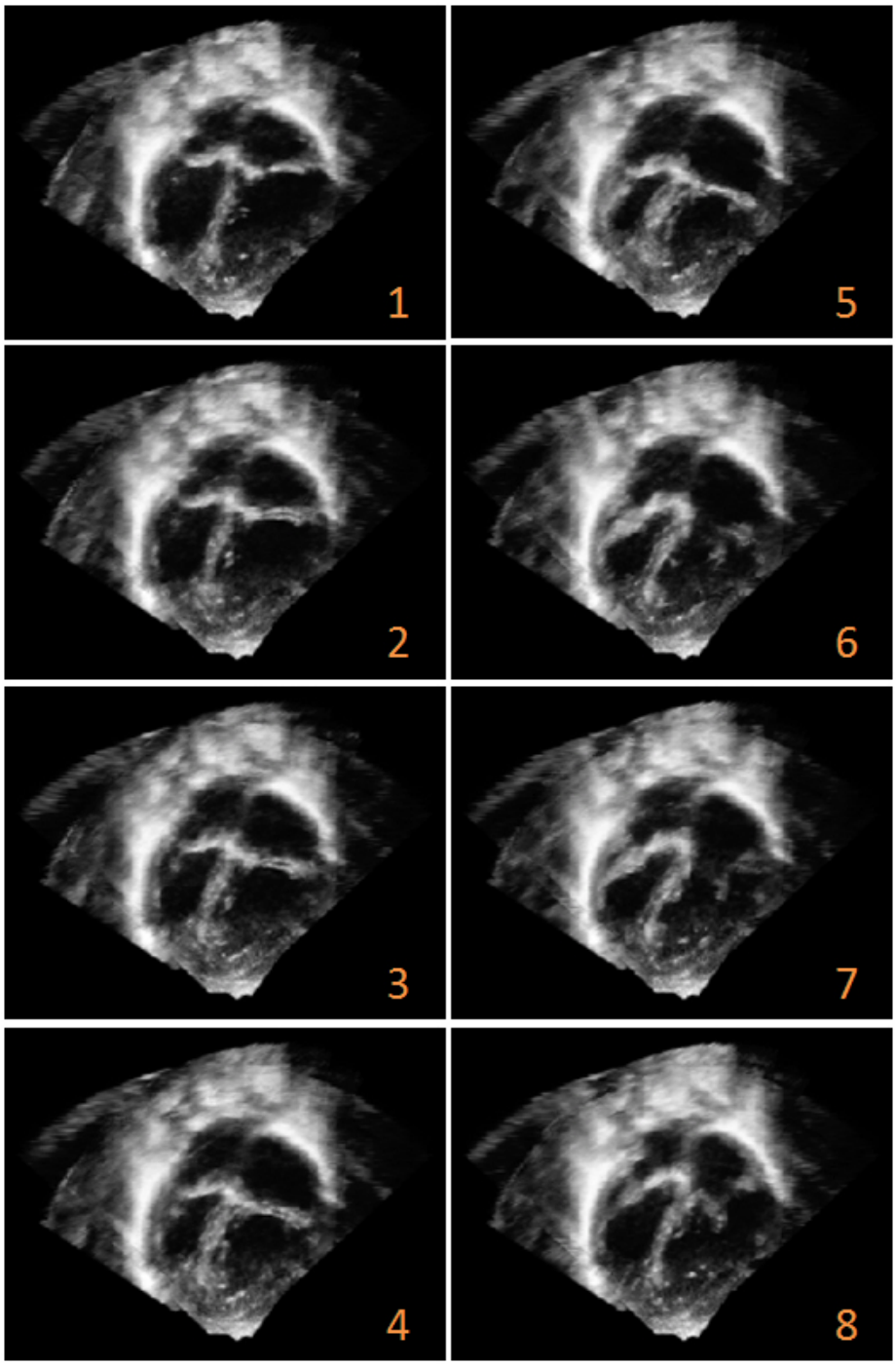


Figure 7.5: Example slices from one cardiac cycle of a 4D compounded image sequence on TGA1 dataset. These images show my proposed 3D compounding technique has been applied to each cardiac phase to produce a visually plausible 4D compounded sequence.

visually plausible 4D compounded sequences. TGA presents unoperated only in children, who normally have a better acoustic window compared to adult patients.

In this section two TGA patients datasets were used as an initial testing sample to see the potential use of my compounding technique using clinical data. Results have shown the compounding method to produce an image which, in particular, showed better border definition. In the next section I apply my algorithm to a larger number of datasets in a different patient group, HLHS.

7.2 Hypoplastic Left Heart Syndrome (HLHS)

7.2.1 Clinical Background

In Hypoplastic left heart syndrome (HLHS) (Noonan and Nadas, 1958), the left side of the heart is severely underdeveloped (see Figure 7.6 for illustration). Most of the anatomical structures on the left side of the heart are too small and underdeveloped (hypoplastic), simply can't function effectively to meet the body's needs for oxygenated blood. Other structures in left heart can also be hypoplastic, such as the mitral vale and aortic valve are normally atretic (blocked).

The right ventricle is positioned very anteriorly under the sternum making it difficult to assess echocardiographically. The shape is complex, making it difficult to estimate from mathematical assumptions based on a single projection. This problem would be overcome if 3D volumes taken from different probe positions could be compounded into a single 3D image with

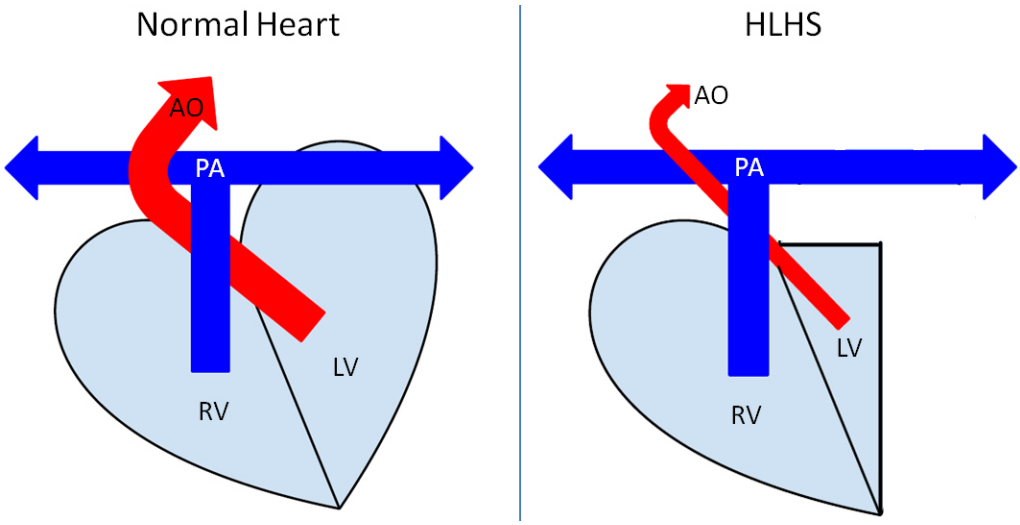


Figure 7.6: Simplified graphes illustrate the difference between normal heart and patient heart with HLHS. In HLSH the left ventricle is smaller and the aorta is narrowed. (LV = left ventricle; RV = right ventricle; PA = pulmonary artery; AO = aorta)

increased field-of-view, as well as improved SNR and better endocardial definition which are good for volume assessment.

7.2.2 Experiments

Five previously acquired datasets (HLHS1-5) from HLHS patients were used with 4 to 12 images per dataset. Typical volumes were $256 \times 272 \times 201$ voxels with $0.38 \times 0.38 \times 0.33\ mm^3$ voxel size.

The transducer was not tracked so manually picked landmarks were again used to acquire the starting estimation. Different points were chosen for these datasets compared to TGA data as the images were primarily of the right ventricle. Six points are picked by an experienced clinician from each of the original images (at the end-diastolic phase) at the following positions: neoaortic valve, left atrium (middle), right atrium (middle), tricuspid valve, apex in right ventricle, and septum between left and right ventricles.

Point based registration was performed using Procrustes alignment (Gower and Dijksterhuis, 2004) and the output transformation used as the starting estimation for my intensity-based registrations RCF (described in Section 5.4.1).

A visual assessment of compounded image quality was carried out. A similar method as described in Chapter 7.1.2 but only two observers were used.

7.2.3 Results

Example image slices are shown in Figure 7.7, 7.8 and 7.9. Both original and compounded images are presented in views which form part of the standard clinical protocol for assessing HLHS patients. These images were shown to an experienced clinician who highlighted the following features: much clearer endocardial border and valves have been seen at 4CH view in Figure 7.7; better visualised borders are shown at RVOT view in Figure 7.8, which is essential for the assessment of function; better boundary definition of the left ventricle is presented at SAX view in Figure 7.8, which is very useful to assess the volume and function of the left ventricle; extended field-of-view (especially greater coverage of right ventricle) and reduced cavity noise have also been seen in compounded image in Figure 7.9, which improves the endocardial border definition and potentially benefit for the assessment of such complex anatomical structure i.e. right ventricle.

The RCF group-wise registration didn't work automatically for all of the data (9 out of 36 volumes failed to register automatically as determined by visual inspection). When failures occurred these were removed from the

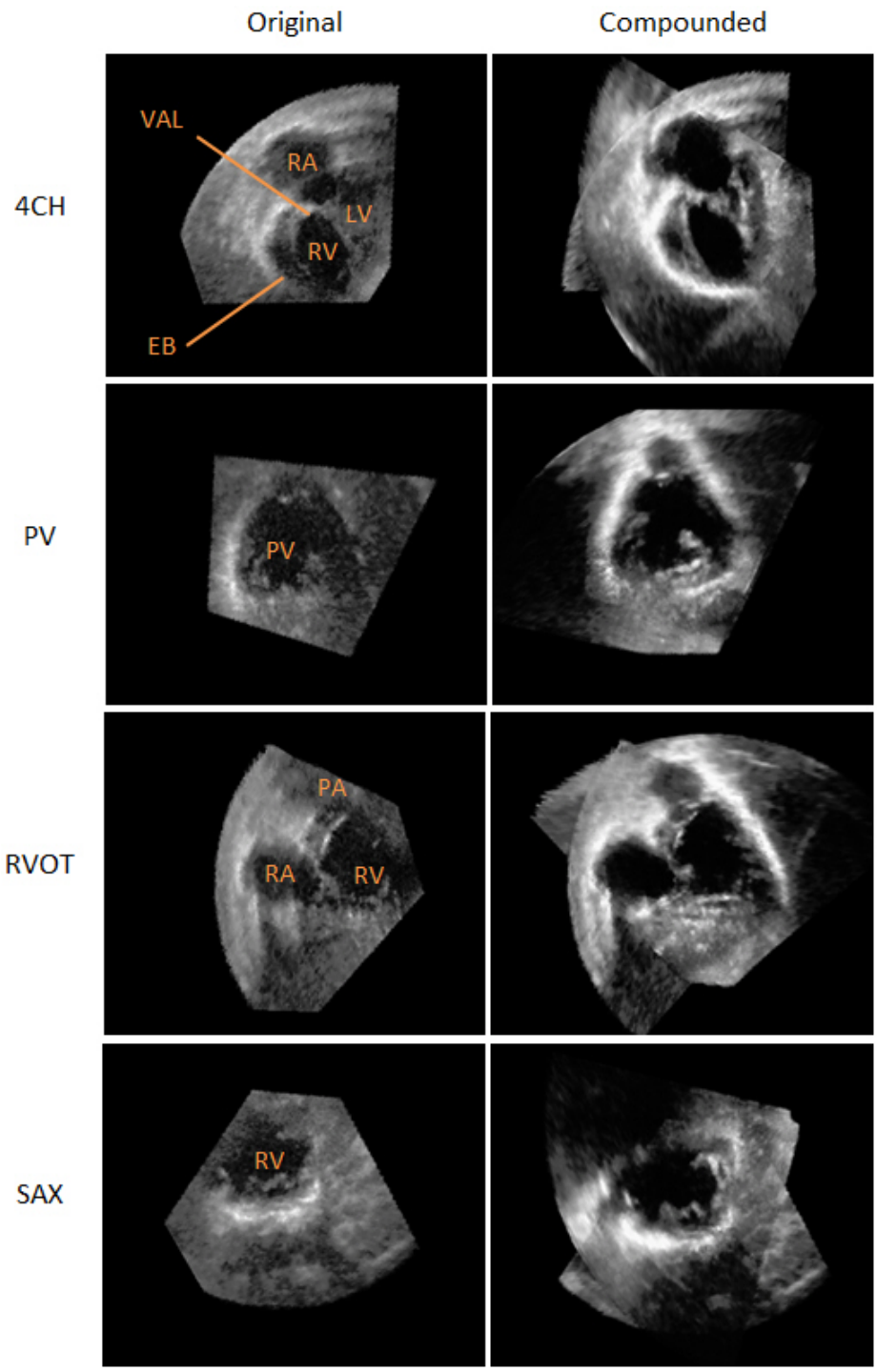


Figure 7.7: Slices from 3D echo images on HLHS1 patient data. Original and compounded images are presented in the following view angles: 4CH (apical four chamber), PV (pulmonary valve), RVOT (right ventricle outflow track) and SAX (short axis). My compounded images shows less noise, more contrast and with RV and PV well presented in single volume. (LV = left ventricle; RV = right ventricle; RA = right atrium; PA = pulmonary artery; EB = endocardial border; VAL = valve)

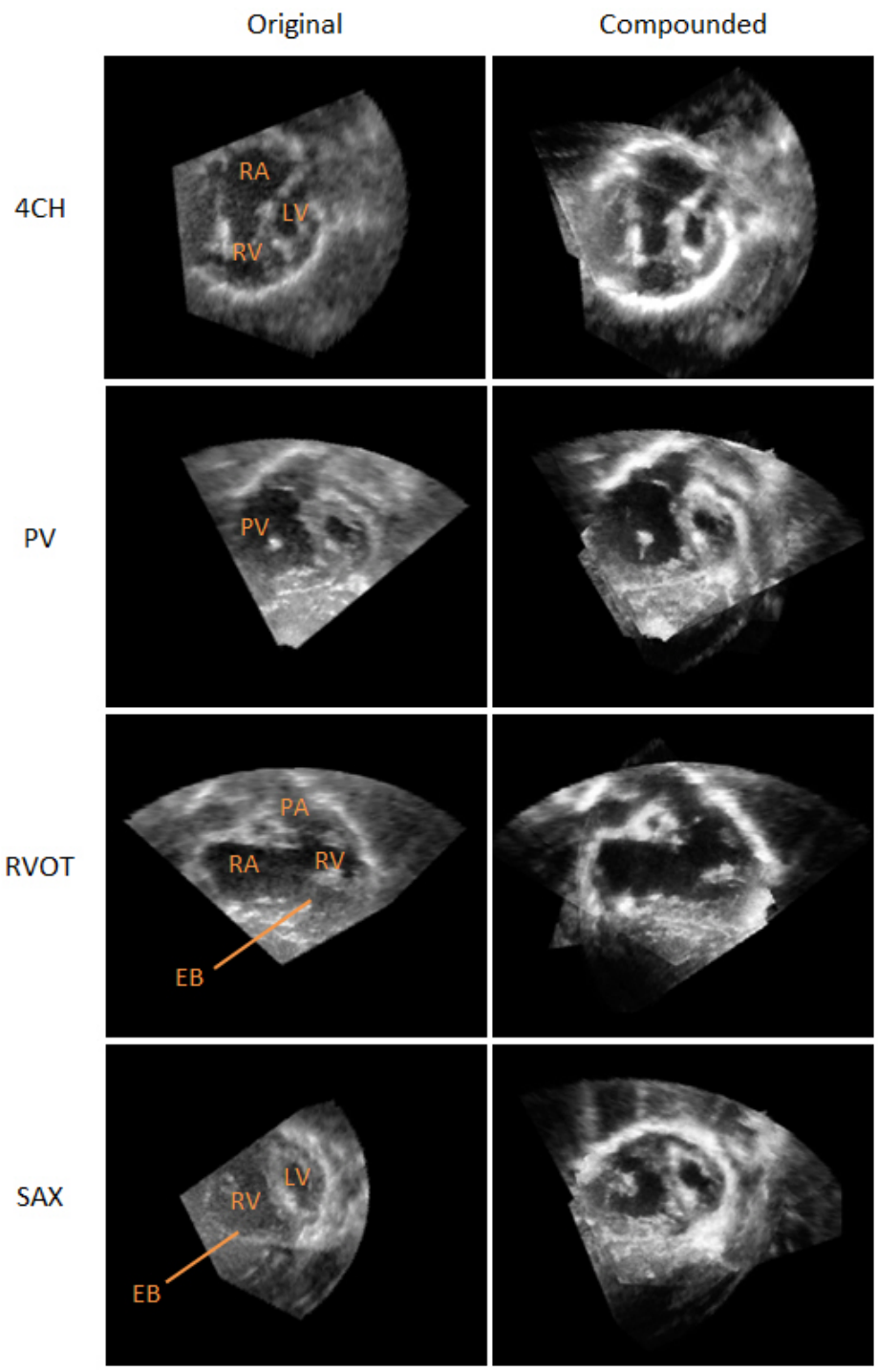


Figure 7.8: Slices from 3D echo images on HLHS4 patient data. Original and compounded images are presented in the following view angles: 4CH (apical four chamber), PV (pulmonary valve), RVOT (right ventricle outflow track) and SAX (short axis). My compounded images shows less noise, more contrast and with RV and PV well presented in single volume. (LV = left ventricle; RV = right ventricle; RA = right atrium; PA = pulmonary artery; EB = endocardial border)

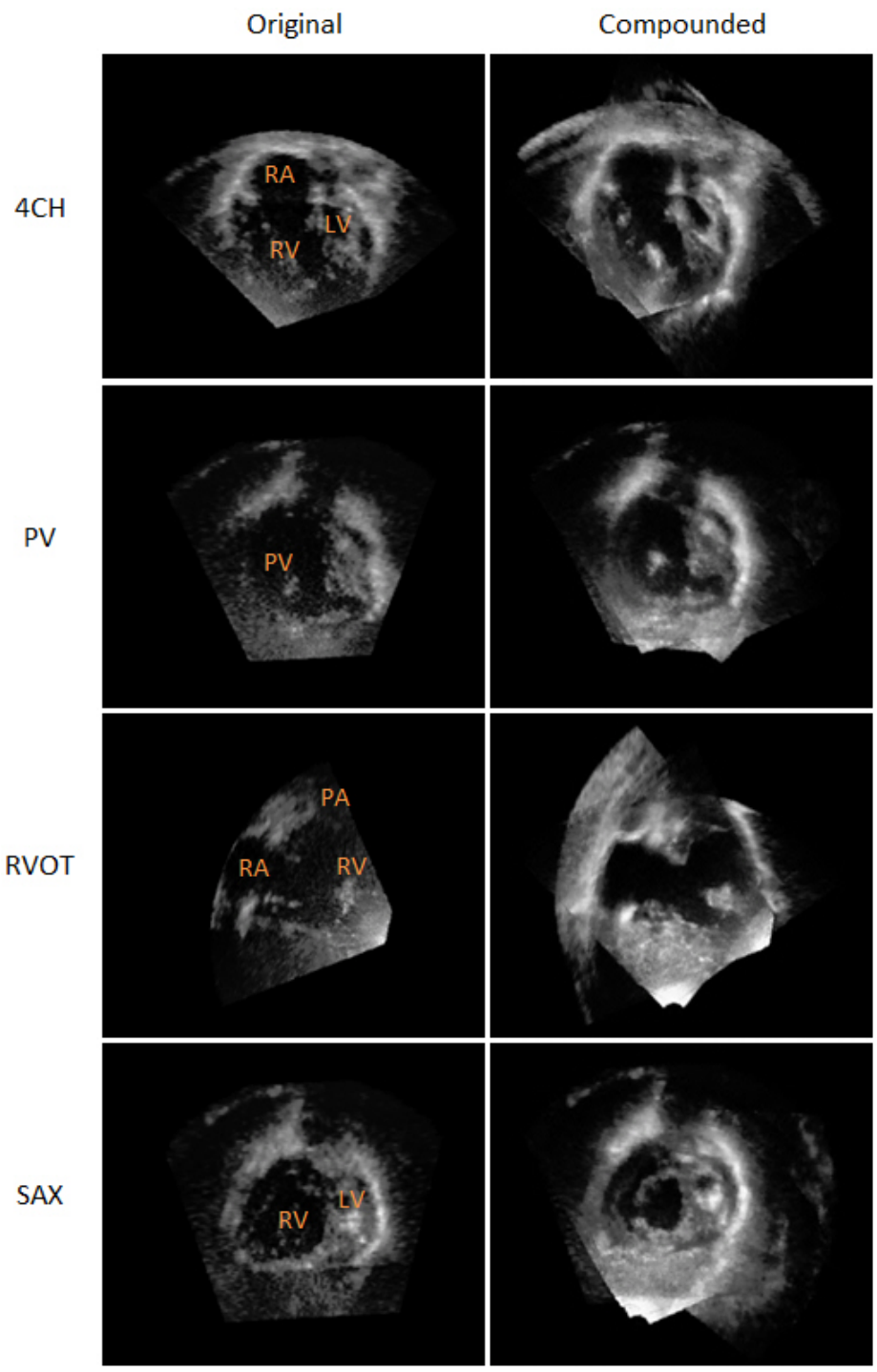


Figure 7.9: Slices from 3D echo images on HLHS5 patient data. Original and compounded images are presented in the following view angles: 4CH (apical four chamber), PV (pulmonary valve), RVOT (right ventricle outflow track) and SAX (short axis). My compounded images shows less noise, more contrast and with RV and PV well presented in single volume. (LV = left ventricle; RV = right ventricle; RA = right atrium; PA = pulmonary artery; EB = endocardial border)

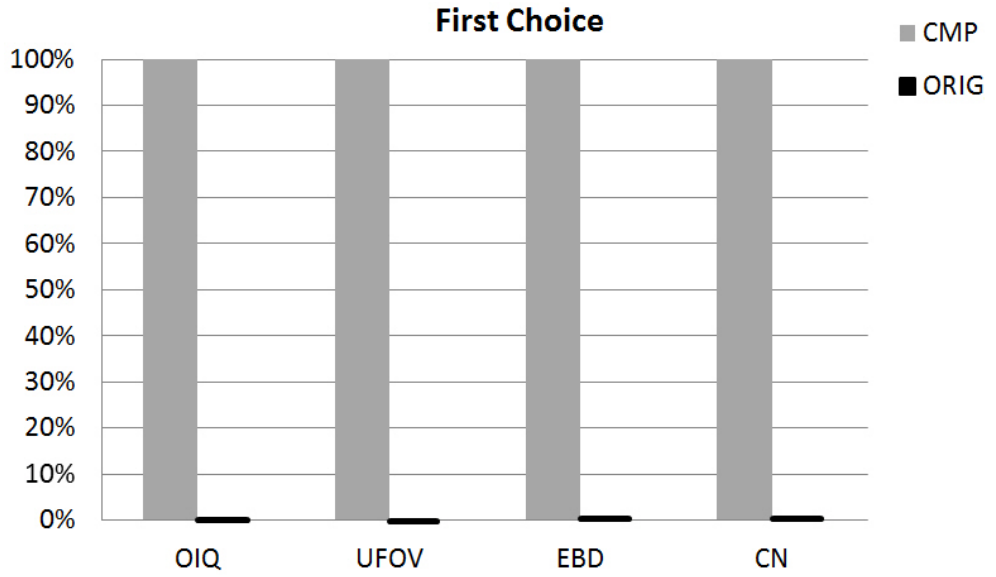


Figure 7.10: Clinical visual inspection result. Percentage of times an image was ranked the best from presented images for each dataset. Presented images were the compounded volume CMP and the original images ORIG. Categories for selection were overall image quality (OIQ), usable field-of-view (UFOV), endocardial border definition (EBD) and cavity noise (CN).

set of data and the compounding process was repeated. In some cases when a large number of failures occurred the images corresponding to the failed registrations were used together to form a separate additional compounded volume.

The clinical visual assessment study results are shown in Figure 7.10. From the five datasets it again shows the clinicians' strong preference for the compounded image rather than individual input images. The compounded image was always chosen as the best image over all four criteria: OIQ, UFOV, EBD and CN.

7.2.4 Discussion

In this section, the compounded volumes show incorporation of all components of the right ventricle with good definition of the endocardial border.

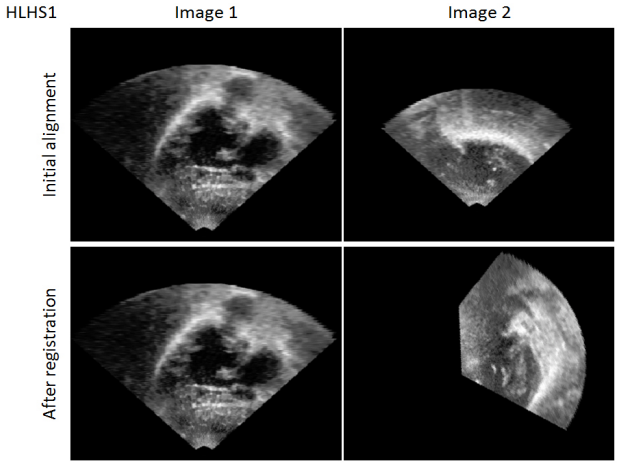
The compounding technique provided enhanced information where a single echocardiographic projection provides incomplete acoustic access to the area of interest. This is particularly useful for the right ventricle which has a much more complex geometric shape than left ventricle.

Registration or compounding tasks proved to be more difficult using this data. The main reason for this is expected to be the starting estimation error. Figure 7.11 shows the situation where registration failed due to the initial alignment error. This is believed to be because it was more difficult to pick corresponding 3D positions on the right ventricle due to its more complex shape and also because overall image quality on the right ventricle was lower. The increased errors in starting position are believed to have resulted in the registration failures for the RCF algorithm.

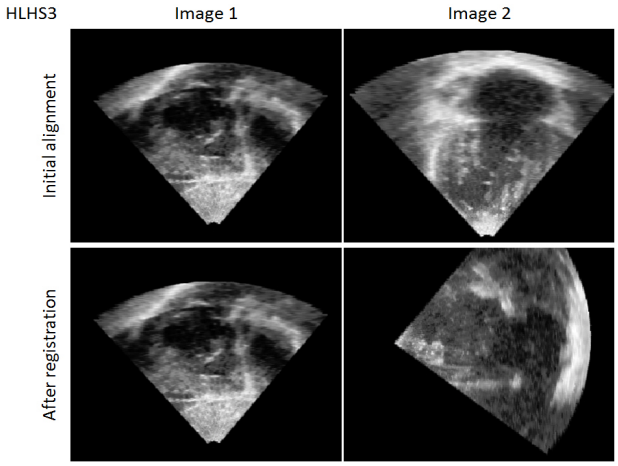
As commented previously use of manual landmarks to produce a starting position is unfeasible in a routine clinical setting. A possible solution to obtaining an sufficiently accurate starting position in an automated way would be to use magnetic tracking. Currently in my institution, we have started to use magnetic tracking devices to track both the transducer and any movement of the patient.

7.3 Conclusion

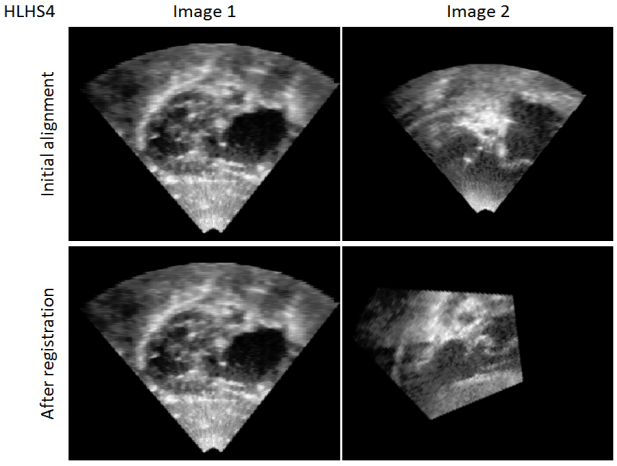
In this chapter, my feature consistency compounding technique has been applied to patient data, where the patient had a congenital heart defect in the form of TGA and HLHS. In total seven datasets have been visually evaluated by experienced clinicians, and in each case the compounded volumes



(a)



(b)



(c)

Figure 7.11: Example images show registration failure due to the initial alignment error. For each case, Image 2 is transformed to the coordinate of Image 1, where the first row presents images with initial alignment which shows the anatomy displayed at different positions or orientations, and the second row shows images after RCF registration where the algorithm can be seen to have failed to register correctly.

provide better visualization of complex anatomy, which is good for assessment and diagnosis for both the TGA and HLHS cases. When accurate registrations were obtained, my compounding algorithm achieved similar performance with patient data compared to volunteer data in the previous chapter.

Further evaluation is planned to see how my compounding technique will improve the volume assessment result of right ventricle. When compared to magnetic resonance imaging, 3D echo produces ventricular volumes which are on average 30 % less than those from MRI (van der Zwaan et al., 2010). Future work regarding to clinical compatibility of my technique and clinical impact will be covered in the next chapter.

This chapter concentrated on an initial clinical evaluation of the visual appearance of my compounded echo. The evidence of how compounded echo image can benefit post-processing such as segmentation and motion estimation, will be discussed in the final chapter of this thesis.

Chapter 8

Conclusions

The main aim of my PhD research was to improve the quality of echo images by combining information from multiple views for applications in both clinical diagnosis and to improve the performance of post-processing algorithms. This chapter will discuss how the work I have done has contributed to this aim. Previous chapters will be summarized, and the novelties and findings of my PhD work will be highlighted. The idea of using echo compounding in post-processing will be demonstrated as my collaborative work with other groups. Future work involving technical improvement and clinical compatibility/impact will be discussed later on.

8.1 Summary of the thesis

Echo is currently widely used in clinical diagnosis and evaluation of many heart conditions. However, as discussed in Chapter 2, echo has disadvantages such as low image quality in terms of noise and artefacts, and limited field-of-view due to the restricted acoustic window. The aim of my PhD

research is to overcome some of these limitations. Therefore, compounding techniques are introduced in Chapter 3 which have the ability of combining multi-view acquisitions to make use of the most useful information from each view to produce a final compounded image with improved image quality and extended field-of-view. There is not a large body of prior work on 3D echo compounding in the literature. The aim of previous work was improve the SNR and contrast of echo images. However, they have only applied their techniques to a few images (most of them used just two images, one used six), and also none of them have directly addressed the effect of artefact on image quality.

Assuming echo images are effected by random noise, then basic signal processing (e.g. signal averaging) can be used to improve SNR. This was investigated by Rohling et al. (1997). The evidence showed that compounding more 2D images should improve SNR. Therefore I started my PhD (as described in Chapter 4) with investigation on compounding more 3D images. By extending the current state-of-the-art phase-based method, it is able to be applied to larger numbers of images and the performance was compared with basic Mean and Maximum methods. Ten volumes were acquired and compounded for each dataset. We limited our acquisition to ten volumes as this was estimated to be near the upper limit of acceptance (due to time constraints) in routine clinical practice. Results showed compounding more 3D images using the extended phase-based method continuously increased SNR and contrast, and as expected Mean methods performed well in terms of SNR and Maximum methods performed well in term of contrast.

I defined three main types of region effects within an echo image: angular

dependent effect on boundary between two tissue types which have different acoustic impedance values; speckle noise in homogenous areas caused by constructive and destructive interference between waves reflected from the numerous small scale structures; and (most commonly low intensity) artefacts caused by strongly reflecting boundaries or poor probe-patient contact. In Chapter 5 I proposed what an ideal compounded echo image should look like, and designed a compounding technique towards attempting to achieve this ideal image. My compounding method filters useful information from multi-view images based on feature consistency. It was designed to directly address the effect of artefacts. By taking advantage of more images, strategies looking for consistent features were used by my compounding technique to detect artefacts, and I proposed artefact features should appear as outliers in a set of registered multi-view images.

Thorough validation has been carried out firstly on a static heart phantom (in Chapter 6). Four sets of phantom images were acquired, some directly from the phantom surface (to have a "gold-standard"), and others by imaging through hard and soft tissue mimicking material to degrade the image quality and generate artefacts. Results showed my method was able to take a set of ten images, degraded by soft and hard tissue artefacts, and produce a compounded image of equivalent quality to images acquired directly from the phantom. Then ten volunteer datasets were used for validation (in Chapter 6). Compared to phantom experiments, the volunteer echo images contain real cardiac motion. Also breathing motion and body movement could affect the accuracy of the initial estimation for the image registration algorithm. My method on phantom and volunteer data achieves almost

the same signal-to-noise improvement as the mean method, while simultaneously almost achieving the same contrast improvement as the maximum method. Both phantom and volunteer results showed that a statistically significant improvement in image quality could be achieved by using an increased number of images (ten compared to five). Visual inspection studies, carried out by three clinicians, showed a very strong preference for my compounded volumes in terms of overall high image quality, large field-of-view, high endocardial border definition and low cavity noise.

Chapter 7 is my initial step to assess how my compounding technique performs using clinical data. My feature consistency compounding technique was applied to two types of clinical cases: Transposition of the Great Arteries (TGA) and Hypoplastic Left Heart Syndrome (HLHS). In total seven datasets have been visually evaluated by experienced clinicians, showing that my compounded volumes provide better visualization of complex anatomy, which is good for assessment and diagnosis for both the TGA and HLHS cases. Moving from volunteer to patient data, there are potential issues with obtaining a starting estimate for image registration, temporal alignment errors with patients who have irregular heart rhythm, and poor quality or limited field-of-view images due to patients' poor acoustic windows. In the HLHS cases, I experienced some registration failures due to the initial alignment error. However, with well registered images, my compounding algorithm still achieved similar performance as the volunteer data.

An overview of my 3D echo compounding work is shown in Table 8.1, allowing direct comparison to the methods detailed in my literature review (Table 3.1).

Table 8.1: An overview of my 3D echo compounding work.

	My 3D echo compounding
Compounding method	Feature consistency based compounding
Registration method	fully automatic optical tracking for starting estimation, followed by an adopted semi-simultaneous strategy with implementation of phase-based similarity measure
Intensity normalisation	linear intensity mapping between images
Directly addressing artefacts	method treats artefacts as outliers thus lower weightings are assigned to those region
Simulated data	0
Phantom data	4: including gold-standard and lower quality (much more clinically realistic) image data
Volunteer data	10
Patient data	7: from two specific clinical cases (HLHS and TGA)
The maximum number of images being compounded	10
Numerical assessment	contrast; SNR; UQI
Visual assessment	clinical visual assessment study

8.2 Future work

8.2.1 Technical improvement

The work described in this thesis is focused on spatial compounding, and so I have compounded end-diastolic images, as the heart is moving slowly, to reduce the chance of temporal misalignment effecting compounded image quality. I have used simple temporal linear interpolation which produced visually plausible 4D compounded sequences for all the data used in this thesis. Errors with using linear interpolation for temporal alignment will probably increase as variations in heart rate increase and for patients with arrhythmia. This was not a problem with our patient datasets as the patient groups used do not have arrhythmia. Further investigation could be carried out for a greater use of ECG. Methods could be devised to detect data if it is too far from a regular rhythm, and then either discard or correct this data. The iE33 echo machine used to acquire all of the images for this thesis has a mechanism to reject wide field-of-view (acquired over four to seven cardiac cycles) 3D images based on comparison of ECG traces, and MR machines also have a similar mechanism to detect and remove data from abnormal rhythms. It is interesting that data from irregular heart cycles may well be of great interest to clinicians and future methods to image these particular heart cycles could have big clinical value.

My compounding method used a rigid intensity based registration algorithm to align the echo volumes. Violations of the rigid body assumption could arise due to deformation caused by breathing motion, or variations in cardiac motion. Registration errors will result in different regions being

compared between echo images. The situation might get worse when larger numbers of images are used and the registration errors could accumulated. In basic compounding methods, such as the mean method, this would cause the blurring of features over a region related to the size of the registration error. However, I believe that my method has within its framework some ability to filter out individual registration errors, as a misregistered image region will have a low consistency value compared to the other images in the dataset and hence a low weighting.

The definition of an "ideal" echo image in Section 5.2 was primarily to define my compounding aims. The specific features in an "ideal" echo image will probably depend on the final use of the image. For example the speckle pattern in homogeneous areas has a random granular appearance and different tissues have different types of speckle pattern as it is formed from scatterers in the tissues. On one hand, speckle degrades both the spatial resolution and contrast in echo images. On the other hand, the appearance of a speckle pattern can be used by clinicians to identify different tissue types, and is also extremely useful for motion tracking. However, speckle will get blurred during the compounding process. Therefore, depending on the clinical application, it may be beneficial to try to preserve speckle. Investigations can be carried out for compounding images while preserving the speckle patterns for motion. This could be done using the same method as Piella et al. (2011) i.e. using the images independently and compound the motion information rather than image information (further details of this method will be given in Section 8.2.3).

8.2.2 Clinical compatibility of my technique

My results in Chapter 6 also show that the use of ten as opposed to five images is significant with respect to numerical measures of image quality and from visual inspection. There will obviously be a limit to the number of volumes that can be acquired as a routine clinical practice. I estimate ten volumes is near to the upper limit of acceptance, and could be obtained in approximately 8 minutes by an experienced clinician following a devised protocol. One method which could allow the acquisition of more images within a clinically acceptable time frame would be to use real time 3D data streaming. This would produce many more images, but they would each have a smaller field-of-view and would likely be of lower quality than the images acquired and used in this submission. This would allow the compounding algorithm to use much more data, but would require the algorithm to have greater ability to register accurately with noisy small field-of-view data.

As commented previously in Chapter 7, use of manual landmarks to produce a starting estimation for image registration algorithm is unfeasible in a routine clinical setting, due to time and space constraints in clinic. Although optical tracking has a high accuracy, it requires a line of sight between transducer and receiver, which could restrict the transducer's motion. Magnetic tracking does not have this line of sight motion restriction for the transducer. However, the accuracy of a magnetic tracking system can be affected by the variation in the external magnetic field, particularly those caused by metal objects. The accuracy of magnetic tracking is very likely to be good enough to provide automatic starting estimates for image

registration.

Current software for registration and compounding are designed for off-line running on CPU (central processing unit) and are fairly computationally expensive. The whole processes take about 15 minutes for registration and 20 minutes for compounding if using 10 images with an average image size of $210 \times 210 \times 210$ voxels. Recently, the use of a GPU (graphics processing unit) to do scientific and engineering computing has become a very popular topic. The sequential part of a software still runs on the CPU, but if the heavily computational part can be modified to a parallel programming model, it will be is accelerated by the GPU. Both the part of the registration algorithm (calculating the similarity measure) and the core part of my compounding algorithm (generating consistency weighting values for all the voxels in all the images) can be modified to parallel fashion, which will result in significant improvement on computational time.

8.2.3 Clinical impact

Echo is a very interactive modality, and the act of compounding takes away part of the interactive real-time nature of an echo examination. Potential areas of benefit include the extended field-of-view which may permit better and more complete imaging of some parts of the heart, such as the right ventricle which are difficult or impossible to incorporate into a single echocardiographic projection. The ability to compound 3D volumes could lead to the production of a single "best" dataset of the heart which could be re-analyzed and post-processed rather than multiple analysis of different 3D volumes. Such a dataset would avoid the real problem of artefacts in

volumes taken from a single projection and could provide easier review of images by clinicians not involved in the acquisition.

In the literature, it has been shown that:

- the ejection fraction and volume-time curves computed using compounded echo image agreed better with MR measurement than those from a single image (Ye et al., 2002);
- the volume measurement results from compounded data were equivalent to those from original data indicating no systematic bias (Szmigielski et al., 2010);
- the performance of left ventricle segmentation and motion tracking improved when using a compounded image compared to a single individual image (Rajpoot et al., 2011).

Future investigations could be carried out on the use of my compounding techniques in post-processing (e.g. cardiac parameter estimation), or its use in associated research programmes (such as: the use of compounded images as an input for computational models, overlay of compounded images onto interventional fluoroscopy during cardiac interventions, or the use of compounded images to aid image registration). During the course of my research, evidence of the use of 3D echo compounding in post-processing has come from my collaborative work with other research groups which are described below.

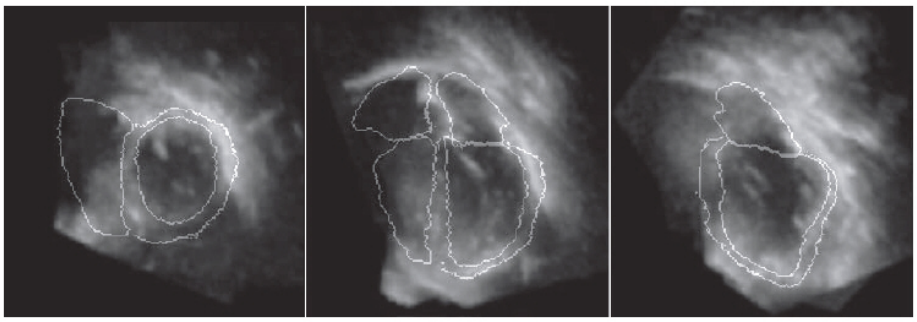
The use of compounding in segmentation

Zhuang et al. (2010) (my collaborative work with University College Lon-

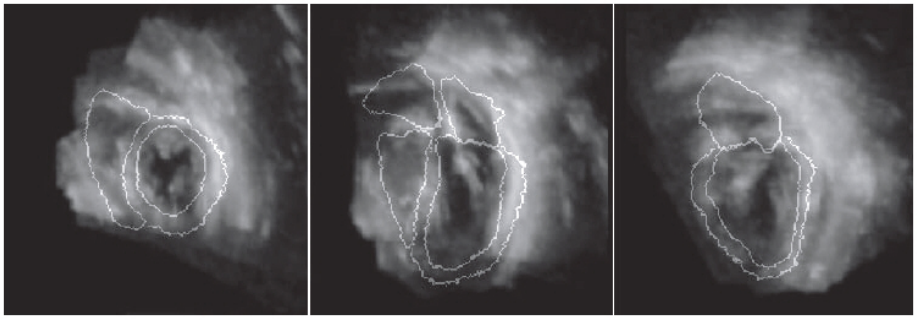
don) presented a framework for automatic whole heart segmentation from 3D compounded echo images. The authors proposed to use the registration-based segmentation framework and adopt a new similarity measure combining local phase, intensity information and local geometry for registration. The experimental results demonstrated the proposed method achieved an accuracy of 6.4 % volume difference against the gold standard for the left ventricle segmentation, comparable to an MR segmentation, and an average accuracy of 14 % for segmentation of all four chambers and myocardium. Figure 8.1 shows the segmentation propagation results.

Automated whole heart segmentation of 3D echo is useful in cardiac functional analysis to achieve quantitative diagnostic information of the heart. However, characteristics of ultrasound imaging such as limited field-of-view, artefacts and inconsistent intensity distribution make automated approaches a challenge. The authors' work is motivated by 3D echo compounding techniques, which can produce compounded 3D echo images with a wider field-of-view and better image quality than a single ultrasound scan. The compounded images provide increased consistency of image intensities between the template and image to be segmented, which resulted in good segmentation accuracy and robustness using intensity information.

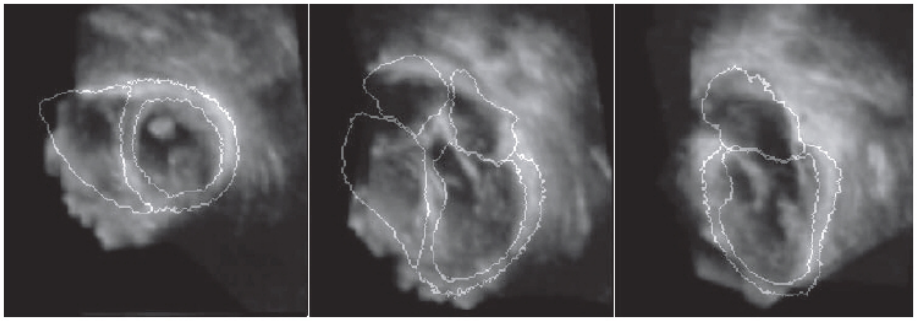
At the time of this collaborative work, only the maximum method was used in compounding a maximum of twelve 3D echo images. As discussed in Chapter 4, the maximum method is good to remove shadowing artefact and extend field-of-view, but SNR stays the same. Also the segmentation performance is affected by the image coverage of the heart. Future work could be to use specific acquisition protocol to improve the coverage and use



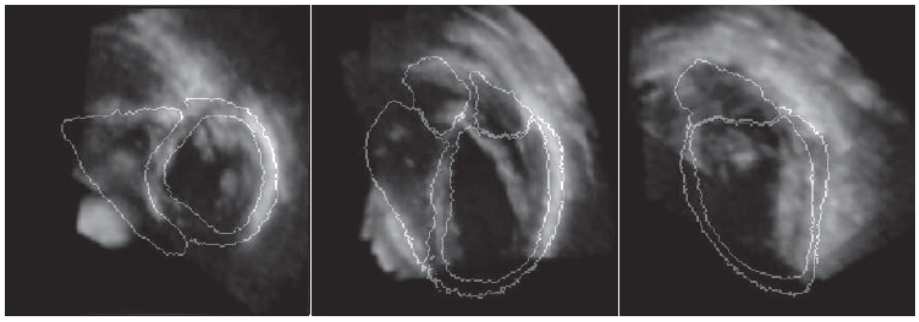
(a)



(b)



(c)



(d)

Figure 8.1: The automated whole heart segmentation propagation results of four cases. Results are displayed in the short-axis view, the four chambers view and the two chambers view. Figure is reproduced from Zhuang et al. (2010).

my feature consistency based compounding for better image quality, which is expected to further improve the automatic whole heart segmentation framework.

The use of compounding in motion estimation

Piella et al. (2011) (my collaborative work with Universitat Pompeu Fabra) presented a new registration framework for estimating myocardial motion and strain from multiple views of 3D+t echo sequences. The authors proposed a multi-view diffeomorphic registration strategy that enforces smoothness and consistency in the spatio-temporal domain by modelling a continuous 3D+t velocity field. It used the original input images, speckle information and image intensity weighting schemes given by 3D echo compounding techniques. Results showed that the inclusion of several views improves the consistency of the strain curves and reduces the number of segments where a non-physiological strain pattern is observed (Figure 8.2).

Strain estimation provides useful information for cardiac functional analysis and the speckle patten information is used as an important feature for motion estimation. However, speckle could be blurred out in the compounding process. The originality of authors' approach is in the estimation of the transformation directly from the multiple views rather than from a single view or a compounded sequence. This allows exploiting all spatio-temporal information available in the input images. However, not all the image information is valuable due to image noise and artefacts. Therefore, 3D compounding techniques are used to generate weighting maps for different view images, and the similarity measure with a weighting scheme balances the

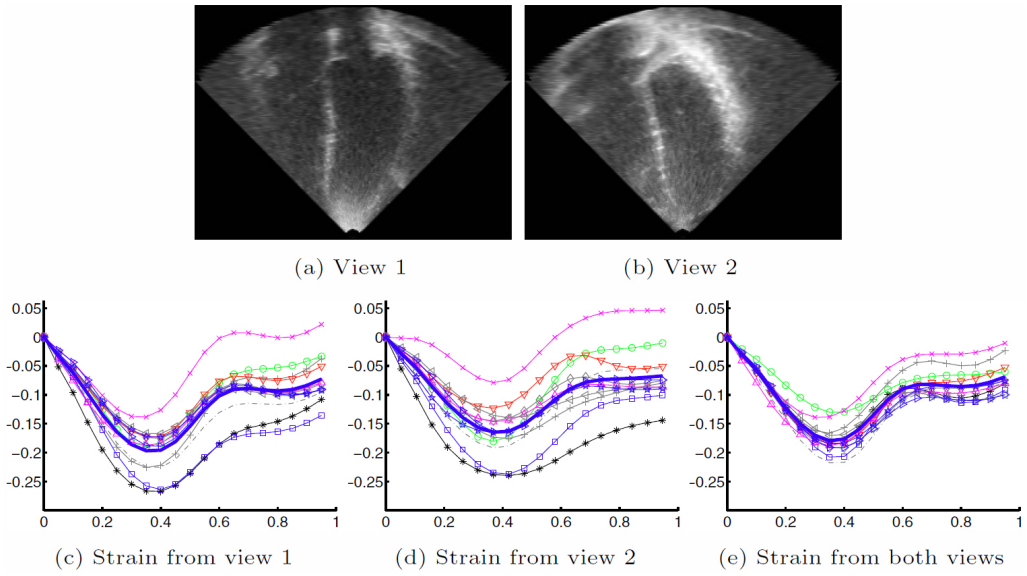


Figure 8.2: Longitudinal strain quantified in basal and mid-level segments. Figure is reproduced from Piella et al. (2011). Full details of the American Heart Association (AHA) segments labelling can be found in the authors' publication.

contribution of the different views in motion estimation process.

This work used an equal weighting scheme in the similarity metric at that time. However, it is ongoing research and more sophisticated weighting information has been generated using my feature consistency based compounding and passed to authors. Further investigation is undergoing.

8.3 Overall summary

In this thesis I have clearly shown that using my consistency based compounding techniques on larger numbers of multi-view 3D echocardiographic images can produce a final compounded image with high quality in terms of reducing effect of artefacts, increase SNR and contrast, and remove angular dependence of boundary reflection. Validation has been carried out using phantom, volunteer and clinical data. Evidence has indicated my com-

pounding could have positive clinical impacts including: benefit for clinical diagnosis, and support for image post-processing tasks. This technique is in the process of being adopted to build a complete system, which will integrate data acquisition, image registration and image compounding all together. It aims to provide a platform for better investigation of congenital heart diseases (e.g. HLHS) and benefit for computational modelling using 3D echocardiography.

Bibliography

Abelmann, W. (1985). Incidence of dilated cardiomyopathy. *Postgraduate Medical Journal*, 61(722).

Aljabar, P. (2008). Tracking longitudinal change using mr image data. *PhD Thesis, Imperial College London*.

Badano, L., Lang, R., and Zamorano, J. (2010). Textbook of real-time three dimensional echocardiography. *Springer*.

Bigatello, L., Haspel, K., Hess, D., Warren, R., and Hurford, W. (2000). Critical care handbook of the massachusetts general hospital. *Lippincott Williams & Wilkins*.

Cooper, J. (2002). Myocarditis: From bench to bedside. *Humana Press*.

Dalvi, R., Hacihaliloglu, I., and Abugharbieh, R. (2010). 3D ultrasound volume stitching using phase symmetry and harris corner detection for orthopaedic applications. *SPIE Medical Imaging*, 762330.

Dendy, P. and Heaton, B. (1999). Physics for diagnostic radiology, second edition. *Bristol and Philadelphia: Institute of Physics Publishing*.

Drukker, K. and Giger, M. L. (2003). Computerized analysis of shadow-

ing on breast ultrasound for improved lesion detection. *Medical Physics*, 30(7):1833–1842.

Duck, F. A. (1990). Physical properties of tissue. *London: Academic*.

Entrekin, R. R., Porter, B. A., Sillesen, H. H., Wong, A. D., Cooperberg, P. L., and Fix, C. H. (2001). Real-time spatial compound imaging: Application to breast, vascular, and musculoskeletal ultrasound. *Seminars in Utrasound, CT, and MRI*, 22(1):50–64.

Felsberg, M. and Sommer, G. (2001). The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136–3144.

Fetics, B., Wong, E., Murabayashi, T., Nelson, G., Cohen, M.-M., Rochitte, C., Weiss, J., Kass, D., and Nevo, E. (2001). Enhancement of contrast echocardiography by image variability analysis. *IEEE Transactions on Medical Imaging (TMI)*, 20(11).

Forsberg, F. (2004). Ultrasonic biomedical technology; marketing versus clinical reality. *Ultrasonics*, 42:17–27.

Frankel, H. and deBoisblanc, B. (2010). Bedside procedures for the intensivist. *Springer*.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association (JASA)*, 32(200).

Gee, A. H., Housden, R. J., Hassenpflug, P., Treece, G. M., and Prager, R. W. (2006). Sensorless freehand 3D ultrasound in real tissue: Speckle

decorrelation without fully developed speckle. *Medical Image Analysis (MedIA)*, 10(2):137–149.

Gower, J. C. and Dijksterhuis, G. B. (2004). Procrustes problems. *Oxford Statistical Science Series*.

Grau, V., Becher, H., and Noble, J. A. (2007). Registration of multiview real-time 3-D echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 26(9):1154–1165.

Grau, V. and Noble, J. A. (2005). Adaptive multiscale ultrasound compounding using phase information. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 8:589–596.

Guo, Y., Werahera, P., Narayanan, R., Li, L., Kumar, D., Crawford, E., and Suri, J. (2009). Image registration accuracy of a 3-dimensional transrectal ultrasound-guided prostate biopsy system. *Journal of Ultrasound in Medicine*, 28(11).

Hill, D. L., Batchelor, P. G., Holden, M., and Hawkes, D. J. (2001). Medical image registration. *Physics in Medicine and Biology (PMB)*, 46(3):R1–R45.

Holden, M., Hill, D., Denton, E., Jarosz, J., Cox, T., Rohlfing, T., Goodey, J., and Hawkes, D. (2000). Voxel similarity measures for 3-d serial mr brain image registration. *IEEE Transactions on Medical Imaging (TMI)*, 19(2).

Hoskins, P. R., Thrush, A., Martin, K., and Whittingham, T. A. (2003).

Diagnostic ultrasound: Physics and equipment. *Greenwich Medical Media Limited*.

Hung, J., Lang, R., Flachskampf, F., Shernan, S. K., McCulloch, M. L., Adams, D. B., Thomas, J., Vannan, M., and Ryan, T. (2007). 3D echocardiography: A review of the current status and future directions. *Journal of the American Society of Echocardiography*, 20(3).

Jefferies, J. and Towbin, J. (2010). Dilated cardiomyopathy. *The Lancet*, 375(9716).

Jensen, J. A. (1996). Field: A program for simulating ultrasound systems. *Medical & Biological Engineering & Computing*, 34:351–353.

Karnik, V., Fenster, A., Bax, J., Cool, D., Gardi, L., Gyacskov, I., Romagnoli, C., and Ward, A. (2010). Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. *Medical Physics*, 37(2).

Kaye, G. W. C. and Laby, T. H. (1995). Tables of physical and chemical constants 16th edn. *London: Longman*.

Kenna, R. and Berche, B. (2011). Normalization of peer-evaluation measures of group research quality across academic disciplines. *Research Evaluation*, 20(2).

Klein, A. and Asher, C. (2011). Clinical echocardiography review: A self-assessment tool. *Lippincott Williams & Wilkins*.

Lang, R. M., Mor-Avi, V., Dent, J. M., and Kramer, C. M. (2009). Three-

dimensional echocardiography: Is it ready for everyday clinical use? *JACC: Cardiovascular imaging*, 2(1).

Law, N. M. (2006). Lucky imaging: Diffraction limited imaging from the ground in the visible. *Cambridge University, UK*.

Ledesma-Carbayo, M., Kybic, J., Desco, M., Santos, A., Suhling, M., Hunziker, P., and Unser, M. (2005). Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. *IEEE Transaction on Medical Imaging (TMI)*, 24(9):1113–1126.

Leotta, D. F. and Martin, R. W. (1999). Three-dimensional spatial compounding of ultrasound scans with incidence angle weighting. *IEEE Ultrasonics Symposium*, 2:1605 –1608.

Ma, Y., Rhode, K. S., Gao, G., King, A. P., Chinchapatnam, P., Schaeffter, T., Hawkes, D. J., Razavi, R., and Penney, G. P. (2008). Ultrasound calibration using intensity-based image registration: for application in cardiac catheterization procedures. *SPIE Medical Imaging*, 6918:69180O–1–69180O–9.

Maintz, J. and Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis (MedIA)*, 2(1):1–36.

Makela, T., Clarysse, P., Sipila, O., Pauna, N., Pham, Q. C., Katila, T., and Magnin, I. (2002). A review of cardiac image registration methods. *IEEE Transactions on Medical Imaging (TMI)*, 21(9):1011–1021.

Martins, P. and Castela, E. (2008). Transposition of the great arteries. *Orphanet Journal of Rare Diseases*, 3(27):1–10.

Maurer, C. R., Fitzpatrick, J. M., and Vanderbilt (1993). A review of medical image registration. In Maciunas, R. J., editor, *Interactive Image-Guided Neurosurgery*, pages 17–44. American Association of Neurological Surgeons, Park Ridge, IL.

Mitra, D., Connolly, D., Jenkins, S., English, P., Birchall, D., Mandel, C., Shrikanth, K., Gregson, B., and Gholkar, A. (2006). Comparison of image quality, diagnostic confidence and interobserver variability in contrast enhanced mr angiography and 2d time of flight angiography in evaluation of carotid stenosis. *The British Journal of Radiology*, 79.

Morrone, M. C. and Owens, R. A. (1987). Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313.

Mrak, M., Grgic, S., and Grgic, M. (2003). Picture quality measures in image compression systems. *EUROCON*, 1.

Noble, J. A. and Boukerroui, D. (2006). Ultrasound image segmentation: A survey. *IEEE Transactions on Medical Imaging*, 25(8):987–1010.

Noonan, J. A. and Nadas, A. S. (1958). The hypoplastic left heart syndrome. *Pediatric Clinics of North America*, 5(4):1029–1056.

Oh, J. K., Seward, J. B., and Tajik, A. J. (2007). The echo manual. *Lippincott Williams & Wilkins*.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):629–639.

Perperidis, A., Cusack, D., McDicken, N., MacGillivary, T., and Anderson, T. (2009). Temporal compounding and its effect on clinical measurements of cardiac ultrasound data. *Medical Image Understanding and Analysis (MIUA)*, pages 17–21.

Perperidis, D., Mohiaddin, R. H., and Rueckert, D. (2005). Spatio-temporal free-form registration of cardiac mr image sequences. *Medical Image Analysis (MedIA)*, 9(5):441–456.

Piella, G., Craene, M. D., Yao, C., Penney, G., and Frangi, A. (2011). Multiview diffeomorphic registration for motion and strain estimation from 3D ultrasound sequences. *Functional Imaging and Modeling of the Heart (FIMH)*, pages 375–383.

Poh, N. (2011). User-specific score normalization and fusion for biometric person recognition. *Advanced Topics in Biometrics*, pages 401–418.

Poon, T. C. and Rohling, R. N. (2006). Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine and Biology (UMB)*, 32(3):357–369.

Rajpoot, K., Grau, V., Noble, J. A., Becher, H., and Szmigielski, C. (2011). The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking. *Medical Image Analysis (MedIA)*, 14(4).

Rajpoot, K., Noble, J. A., Grau, V., Szmigielski, C., and Becher, H. (2009). Multiview RT3D echocardiography image fusion. *Functional Imaging and Modeling of the Heart (FIMH)*, pages 134–143.

Rakar, S., Sinagra, G., Di Lenarda, A., Poletti, A., Bussani, R., Silvestri, F., and Camerini, F. (1997). Epidemiology of dilated cardiomyopathy. a prospective post-mortem study of 5252 necropsies. *European Heart Journal*, 18(1).

Rhode, K. S., Hill, D. L. G., Edwards, P. J., Hipwell, J., Rueckert, D., Sanchez-Ortiz, G., Hegde, S., Rahunathan, V., and Razavi, R. (2003). Registration and tracking to integrate x-ray and mr images in an xmr facility. *IEEE Transactions on Medical Imaging*, 22(11):1369–1378.

Rohling, R., Gee, A., and Berman, L. (1997). Three-dimensional spatial compounding of ultrasound images. *Medical Image Analysis*, 1(3):177–193.

Rua, E., Castro, J., and Mateo, C. (2008). Quality-based score normalization for audiovisual person authentication. *International Conference on Image Analysis and Recognition (ICIAR)*, 5112.

Schmidt, J., Berg, D., Ploeg, H., and Ploeg, L. (2009). Precision, repeatability and accuracy of optotrak optical motion tracking systems. *International Journal of Experimental and Computational Biomechanics (IJECB)*, 1(1).

Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J., and Miyasaka, K. (2002). Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *International Journal of Experimental and Computational Biomechanics (IJECB)*, 53(4).

Shechter, G., Ozturk, C., Resar, J., and McVeigh, E. (2004). Respiratory motion of the heart from free breathing coronary angiograms. *International Journal of Experimental and Computational Biomechanics (IJECB)*, 23(8).

Shekhar, R., Zagrodsky, V., Garcia, M., and Thomas, J. (2004). Registration of real-time 3-D ultrasound images of the heart for novel 3-D stress echocardiography. *IEEE Transaction on Medical Imaging (TMI)*, 23(9):1141–1149.

Soler, P., Gerard, O., Allain, P., Saloux, E., Angelini, E. D., and Bloch, I. (2005a). Comparison of fusion techniques for 3D+T echocardiography acquisition from different acoustic windows. *Computers in Cardiology*, pages 141–144.

Soler, P., Villain, N., Bloch, I., and Angelini, E. D. (2005b). Volume reconstraction of breast echography from anisotropically degraded scans.

Sonke, J., Lebesque, J., and van Herk, M. (2008). Variability of four-dimensional computed tomography patient models. *International Journal of Experimental and Computational Biomechanics (IJECB)*, 70(2).

Szmigielski, C., Rajpoot, K., Grau, V., Myerson, S. G., Holloway, C., Noble, J. A., Kerber, R., and Becher, H. (2010). Real-time 3D fusion echocardiography. *JACC: Cardiovascular imaging*, 3(7).

Tavazzi, L. (1997). Epidemiology of dilated cardiomyopathy: a still undetermined entity. *European Heart Journal*, 18(1).

Thavendiranathan, P., Liu, S., Datta, S., Walls, M., Nitinunu, A.,

Van Houten, T., Tomson, N., Vidmar, L., Georgescu, B., Wang, Y., Srinivasan, S., De Michelis, N., Raman, S., Ryan, T., and Vannan, M. (2012). Automated quantification of mitral inflow and aortic outflow stroke volumes by three-dimensional real-time volume color-flow doppler transthoracic echocardiography: comparison with pulsed-wave doppler and cardiac magnetic resonance imaging. *Journal of the American Society of Echocardiography*, 25(1).

Towbin, J., Lowe, A., Colan, S., Sleeper, L., Orav, E., Clunie, S., Messere, J., Cox, G., Lurie, P., Hsu, D., Canter, C., Wilkinson, J., and Lipshultz, S. (2006). Incidence, causes, and outcomes of dilated cardiomyopathy in children. *European Heart Journal*, 296(15).

Treece, G. M., Prager, R. W., Gee, A. H., and Berman, L. (2002). Correction of probe pressure artifacts in freehand 3D ultrasound. *Medical Image Analysis (MedIA)*, 6(3):199–214.

van de Kraats, E., Penney, G., Tomazevic, D., van Walsum, T., and Niessen, W. (2005). Standardized evaluation methodology for 2-d-3-d registration. *IEEE Transactions on Medical Imaging (TMI)*, 24(9).

van den Elsen, P. A., Pol, E.-J. D., and Viergever, M. A. (1993). Medical image matching - a review with classification. *IEEE Engineering in Medicin and Biology*, 12(1):26–39.

van der Zwaan, H. B., Helbing, W. A., McGhie, J. S., Geleijnse, M. L., Luijnenburg, S. E., Roos-Hesselink, J. W., and Meijboom, F. J. (2010). Clinical value of real-time three-dimensional echocardiography for right

ventricular quantification in congenital heart disease: Validation with cardiac magnetic resonance imaging. *Journal of the American Society of Echocardiography*, 23(2):134–140.

Wachinger, C. (2007). Three-dimensional ultrasound mosaicing. *PhD Thesis, Technical University of Munich.*

Wachinger, C., Klein, T., and Navab, N. (2012). Locally adaptive nakagami-based ultrasound similarity measures. *Ultrasonics*, 52(4).

Wachinger, C. and Navab, N. (2009). Similarity metrics and efficient optimization for simultaneous registration. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 667 –674.

Wachinger, C., Wein, W., and Navab, N. (2007). Three-dimensional ultrasound mosaicing. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 327–335.

Walsh, C., Dowling, A., Meade, A., and Malone, J. (2005). Subjective and objective measures of image quality in digital fluoroscopy. *Radiation Protection Dosimetry*, 117(1-3).

Wang, Z. and Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Processing Letters*, 9(Y):81–84.

Whatmough, C., Guitian, J., Baines, E., Benigni, L., Mahoney, P. N., Mantis, P., and Lamb, C. R. (2007). Ultrasound image compounding: effect on perceived image quality. *Veterinary Radiology & Ultrasound*, 48(2):141–145.

Whiting, R. B. and Wiens, R. D. (1977). Clinical use of echocardiography: A review. *Southern Medical Journal*, 70(1).

Ye, X., Noble, J. A., and Atkinson, D. (2002). 3-D freehand echocardiography for automatic left ventricle reconstruction and analysis based on multiple acoustic windows. *IEEE Transactions on Medical Imaging (TMI)*, 21(9).

Zhang, W., Noble, J., and Brady, J. (2007). Adaptive non-rigid registration of real time 3D ultrasound to cardiovascular mr images. *Information Processing in Medical Imaging (IPMI)*, 20.

Zhuang, X., Yao, C., Ma, Y., Hawkes, D., Penney, G. P., and Ourselin, S. (2010). Registration-based propagation for whole heart segmentation from compounded 3D echocardiography. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1093–1096.

Zitova, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000.

Publications and Awards

Journal Papers

- Yao, C., Simpson, J. M., Schaeffter, T., and Penney, G. P., 3D Echocardiography Compounding using Large Numbers of Multi-view Acquisitions Based on Feature Consistency, *Physics in Medicine and Biology (PMB)*, vol. 56, pp. 6109-6128, 2011. (refer to thesis Chapter 5)

- King, A. P., Rhode, K. S., Ma, Y., Yao, C., Jansen, C. H. P., Razavi, R., and Penney, G. P., Registration preprocedure volumetric images with intraprocedure 3-D ultrasound using an ultrasound imaging model, *IEEE Transactions on Medical Imaging (TMI)*, vol. 29, pp. 924-937, 2009. (refer to thesis Chapter 5)

Conference Papers

- Piella, G., Craene, M. D., Yao, C., Penney, G. P., and Frangi, A. F., Multiview Diffeomorphic Registration for Motion and Strain Estimation from 3D Ultrasound Sequences, *Functional Imaging and Modeling of the Heart (FIMH)*, vol. 6666, pp. 375-383, 2011. (refer to thesis

Chapter 8)

- Gomez, A., Simpson, J. M., Yao, C., Schaeffter, T., and Penney, G. P., 3D Flow Reconstruction from Multiple Registered Echo Doppler Views, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 879-882, 2011. (refer to thesis Chapter 4)

- Gomez, A., Yao, C., Simpson, J. M., Schaeffter, T., and Penney, G. P., Reconstruction of 3D Flow from Multiple Echo Doppler Views, *Medical Image Understanding and Analysis (MIUA)*, pp. 39-43, 2010. (refer to thesis Chapter 4)

- Yao, C., Simpson, J. M., Schaeffter, T., and Penney, G. P., Spatial compounding of large numbers of multi-view 3D echocardiography images using feature consistency, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 968-971, 2010. (refer to thesis Chapter 5)

- Zhuang, X., Yao, C., Ma, Y., Hawkes, D., Penney, G. P., and Ourselin, S., Registration-based Propagation for whole heart segmentation from compounded 3D echocardiography, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 1093-1096, 2010. (refer to thesis Chapter 8)

- Ma, Y., Penney, G. P., Bos, D., Frissen, P., Fockert, G. D., King, A. P., Gao, G., Yao, C., Totman, J., Ginks, M., Rinaldi, C., Razavi, R., and Rhode, K. S., Evaluation of a robotic arm for echocardiography to X-ray image registration during cardiac catheterization procedures,

IEEE Engineering in Medicine and Biology Society (EMBC), vol. 1, pp. 5829-5832, 2009. (refer to thesis Chapter 4)

- Ma, Y., Penney, G. P., Bos, D., Frissen, P., Fockert, G. D., Yao, C., King, A. P., Gao, G., Rinaldi, C. A., Razavi, R., and Rhode, K., Using a Robotic Arm for Echocardiography to X-ray Image Registration during Cardiac Catheterization Procedures, *MICCAI Workshop on Cardiovascular Interventional Imaging and Biophysical Modelling*, vol. CI2BM09, 2009. (refer to thesis Chapter 4)

- King, A. P., Ma, Y., Yao, C., Jansen, C. H. P., Razavi, R., Rhode, K. S., and Penney, G. P., Image-to-Physical Registration for Image-Guided Interventions Using 3-D Ultrasound and an Ultrasound Imaging Model, *Information Processing in Medical Imaging (IPMI)*, vol. 5636, pp. 188-201, 2009. (refer to thesis Chapter 3)

- Yao, C., Simpson, J. M., Jansen, C. H. P., King, A. P., and Penney, G. P., Spatial compounding of large sets of 3D echocardiography images, *SPIE Medical Imaging*, vol. 7265, pp. 726515-1-8, 2009. (refer to thesis Chapter 4)

- Yao, C. and Penney, G. P., Spatial compounding of 3D echocardiography: novel methodologies for large sets of images, *Medical Image Understanding and Analysis (MIUA)*, pp. 69-73, 2008. (refer to thesis Chapter 4)

Awards

- Susan Tucker Memorial Prize - *Guy's and St Thomas' Hospital, London*, 2009
- Dorothy Hodgkin Postgraduate Awards (DHPA) - *Medical Research Council (MRC) and Philips Healthcare*, 2008
- Overseas Research Students Awards (ORS) - *King's College London*, 2008
- Best Paper Presentation Prize - *Medical Image Understanding and Analysis (MIUA), Dundee*, 2008